UNIVERSITÉ DU QUÉBEC À MONTRÉAL

THE USE OF SINGULAR VECTORS IN THE STUDY OF CANADIAN REGIONAL CLIMATE MODEL INTERNAL VARIABILITY

DISSERTATION

PRESENTED

AS PARTIAL REQUIREMENT

FOR PHD DEGREE IN ENVIRONMENTAL SCIENCES

BY

EMILIA PAULA DIACONESCU

AUGUST 2011

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

APPLICATION DES VECTEURS SINGULIERS DANS L'ÉTUDE DE LA VARIABILITÉ INTERNE D'UN MODÈLE RÉGIONAL PILOTÉ

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN SCIENCES DE L'ENVIRONNEMENT

PAR

EMILIA PAULA DIACONESCU

AOÛT 2011

REMERCIEMENTS

Je tiens à remercier sincèrement toutes les personnes qui m'ont aidée par leur assistance, patience et encouragement durant ces quatre ans d'études doctorales.

Premièrement, je veux exprimer toute ma gratitude à mon directeur, le professeur René Laprise, qui a été le moteur guidant mon développement professionnel. Ces quelques lignes ne suffiront pas à traduire toute la reconnaissance que je lui porte. Toujours optimiste, il m'a corrigée, poussée et encouragée à résoudre les difficultés du moment. Je le remercie chaleureusement pour ses conseils et sa permanente disponibilité. Je veux aussi remercier mon codirecteur, le professeur associé Ayrton Zadra, pour les nombreuses réunions et ses recommandations qui ont enrichi mes réflexions tout au long de mon cheminement.

Je tiens à remercier également les membres de mon comité de thèse, Messieurs Jean-Pierre Blanchet et Ramón de Elía, pour avoir contribué à l'encadrement de ce projet doctoral. Je remercie également les membres du jury qui ont accepté de réviser ma thèse. Merci à la Pr. Eugenia Kalnay, au Pr. Pierre Gauthier et au Dr. Bernard Dugas pour avoir accepté d'évaluer mon travail de thèse. Leurs remarques et recommandations m'ont permis d'apporter des améliorations à la version finale de la thèse.

Les trois premières années de cette étude ont été financées par le Conseil de recherches en sciences naturelles et en génies du Canada (CRSNG). Ce financement a été hautement apprécié et je tiens à le remercier. Le Réseau canadien en modélisation et diagnostics du climat régional (MDCR/ESCER/UQAM) a gracieusement complété le financement pendant la dernière période de recherche et fourni les équipements informatiques. Ma reconnaissance va aussi à toute l'équipe du Réseau MDCR : professeurs, collègues, personnel administratif et technique. Un remerciement particulier à Michel Giguère pour son aide à mes débuts informatiques et à Leo Šeparović et Katja Winger pour avoir éclairci mon chemin pendant mes premières simulations avec MRCC5. Je tiens à remercier aussi Georges Huard, Mourad Labassi et Nadjet Labassi pour le soutien informatique indispensable dans ce travail.

Finalement, ma grande reconnaissance va vers mon conjoint Gabriel Meloche pour l'encouragement, l'amour et l'équilibre apporté dans ma vie, à mes chers parents et mon frère qui, malgré la distance, ont toujours été à côté de moi pour me soutenir et m'encourager. Une mention spéciale à Patrice Meloche pour la révision des textes en anglais. Je tiens à remercier aussi mes amies, Dorina Surcel et Cristina Lupu, pour leur soutien moral et les bons moments passés ensemble au travail et en dehors. Un grand merci à Cristina Lupu qui m'a convaincue de venir au Canada pour continuer mes études à l'Université du Québec à Montréal.

TABLE DES MATIERES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES SIGLES ET ACRONYMES

3D-VAR	Three-dimensional variational data assimilation scheme
4D-VAR	Four-dimensional variational data assimilation scheme
ADM	Adjoint model
AGCM	Atmospheric general circulation models
AMIP2	Atmosphere Model Intercomparison Project
BoM	Bureau of Meteorology
CFCAS	Canadian Foundation for Climate and Atmospheric Sciences
CI	Conditions initiales
CRCM	Canadian Regional Climate Model
CRCMD	Canadian Regional Climate Modelling and Diagnostic Network
DOTSTAR	Dropwindsonde Observation for Typhoon Surveillance near the Taiwan Region
ECMWF	European Centre for Medium-Range Weather Forecasts
ECMWF IFS	ECMWF Integrated Forecasting System
EDA	Ensemble Data Assimilation
EM	Ensemble mean
ENSO	El Niño–Southern Oscillation
EOF	Empirical orthogonal function
EPS	Ensemble Predictions System
ERA40	ECMWF reanalysis
FASTEX	Fronts and Atlantic Storm-Track Experiment
GEM	Global Environnemental Multiscale

ICs	Initial Conditions
ISBA	Model land-surface scheme
IV	Internal variability
JMA	Japan Meteorological Agency
MCG	Modèles atmosphériques de circulation générale
MAMS2	Mesoscale Adjoint Modeling System
MLB	Model Lateral Boundaries
MRC	Modèle régional du climat
MRCC	Modèle régional canadien du climat
MSC	Meteorological Service of Canada
NCAR	National Center for Atmospheric Research
NCAR MM5	NCAR Mesoscale Model
NCEP	National Centers for Environmental Prediction
NORPEX	North Pacific Experiment
NSERC CGSD3	Natural Sciences and Engineering Research Council of Canada Graduate Scholarship
NWP	Numerical Weather Prediction
OTI	Optimization time interval
pdf	Probability density function
PV	Potential vorticity
RCM	Regional Climate Model
SMC	Service Météorologique du Canada
SST	Sea surface temperature
SVs	Singular Vectors

T21, T42, T63	Horizontal triangular spectral truncations
TC	Tropical cyclone
TEPS	Typhoon Ensemble Prediction System
TL63, TL159	Horizontal triangular spectral truncations with linear Gaussian grid
TLM	Tangent linear model
TORPEX PARC	Multinational Observing System Research and Predictability Experiment for Pacific Asian Region Campaign
VI	Variabilité interne
VS	Vecteur singulier
WEPS	One-week Ensemble Prediction System
WSRP	Winter Storm Reconnaissance Program

LISTE DES SYMBOLES

$\circ$	Degré
$(.)^*$	Opérateur adjoint
$(.)^T$	Opérateur de transposition
$(.)_0$	Indice relatif au moment initial
$(.)_h$	Indice relatif à une quantité intégré sur la verticale
$(.)_{ref}$	Indice relatif à la simulation de référence
$(.)_t$	Indice relatif au moment final
$\langle .,. \rangle$	Produit scalaire
$\|\|\|_E$	Norme E
%	Pourcent
∇	Opérateur gradient d'une fonction
∂	Opérateur de différentiation partielle
A	Domaine horizontal du modèle
A_F	Jacobiene de F
A_{free}	Domaine horizontal d'analyse libre du forçage des frontières latérales
c_p	Chaleur spécifique à pression constante de l'air
D	Divergence
E	Énergie des perturbations
F(X)	Équation de tendence du modèle non-linéaire
f	Fonction de location pour le ciblage des observations

G	Sous-espace du système échantillonné par les vecteurs singuliers
g	Constante gravitationnelle
H	Modèle non-linéaire
L	Modèle tangent linéaire propagateur
L_c	Chaleur latente de condensation par l'unité de masse
L_s	Modèle tangent linéaire propagateur normalisé
M	Nombre de simulations MRCC
N	Nombre de vecteurs singuliers
P	Opérateur de projection locale
p	Index de projection
p_r	Pression de référence
p_s	Pression de surface
p_{top}	Pression au sommet du modèle
q	Rapport de mélange
R	Constante des gaz parfaits pour air sec
r	Taux de croissance en six heures
S_{int}	Sources et puits de variabilité interne situés à l'intérieur du domaine
S_{MLB}	Puits de variabilité interne causés par le forçage des frontières latérales
T	Température
T_r	Température de référence
u	Vent zonal
v	Vent meridional
V	Vecteur vent horizontal

Vol	Volume d'intégration
W	Matrice de vecteurs singuliers de gauche
w	Vecteur singulier de gauche
X	Vecteur de l'état de l'atmosphère
x	Perturbation MRCC
Δx	Partie non-projetée de la perturbation x
$\tilde{x}$	Partie projetée de la perturbation x
Y	Matrice de vecteurs singuliers de doite
y	Vecteur singulier de droite
$\hat{y}$	Vecteur singulier avec une norme unitaire
α	Coefficient de projection
ε	Constante de pondération pour le terme d'humidité
η	Coordonnée verticale du modèle
Λ	Matrice de valeurs singuliers
λ	Valeur singulier
λ^2	Facteur d'amplification
ζ	Tourbillon

RÉSUMÉ

La variabilité interne (VI) est une propriété des systèmes chaotiques qui fait que des simulations numériques lancées de conditions initiales différant même de façon minime entre elles, vont éventuellement diverger dans le temps. Des études antérieures ont montré que la VI d'un modèle régional du climat (MRC) avait un caractère épisodique dépendant des conditions synoptiques du moment. Ce projet a pour but principal de vérifier l'hypothèse selon laquelle les maxima notés dans la variation temporelle de la VI sont dus à la croissance rapide des perturbations développées dans des régions atmosphériques dynamiquement instables. L'élément déclencheur est représenté par des instabilités hydrodynamiques résultant de conversions barocline ou barotrope d'énergie. Pour décrire l'espace instable des perturbations, nous avons fait appel à la technique des vecteurs singuliers (VS). Un ensemble de 21 simulations qui diffèrent seulement dans les conditions initiales a été réalisé sur le continent nord-américain et la VI a été exprimée en termes d'énergie totale des perturbations par rapport à une simulation de référence. Plusieurs séries de VS ont été calculées pour trouver les perturbations avec la plus rapide croissance linéaire par rapport à la norme de l'énergie totale sur une période de 36 heures.

L'analyse de la variation totale en 36 heures a montré que la croissance associée au maximum principal de VI était expliquée dans une proportion de 73% par les dix premiers VS et dans une proportion de 51% par le "premier" VS, c'est à dire celui croissant le plus rapidement. Les VS ainsi identifiés avaient des structures de petite échelle spatiale, de forts taux de croissance dans le temps, et ils se développaient à l'intérieur d'un environnement constitué de perturbations matures avec des faibles taux de croissance. Pour un épisode de forte VI, une très grande ressemblance a été trouvée entre la structure des perturbations dans les simulations du MRCC et le VS avec la croissance la plus rapide après 24 à 36 heures d'intégration du modèle linéaire tangent. Au temps initial, le premier VS avait une structure verticale inclinée vers ouest et l'énergie totale était dominée par la composante de l'énergie potentielle. Au temps final, cette inclinaison avait beaucoup diminué et l'énergie totale était alors dominée par l'énergie cinétique, indiquant que la conversion barocline représentait le processus dominant dans l'augmentation de la VI pour cette période. Pour d'autres périodes dominées par les sources d'instabilité, le pourcentage de la variation totale en 36 heures expliqué par les 4 ou 5 premiers VS variait entre 36% et 85%. Toutefois, dans ces cas, le pourcentage n'était pas dominé par le premier VS. Nous avons montré également que dans les périodes dominées par le transfert de VI à l'extérieur du domaine, la projection sur les VS était très petite malgré la présence de sources de VI à l'intérieur du domaine d'analyse. Le fait que ces sources de VI ne soient pas représentées par les plus rapides VS indique qu'elles peuvent être causées par des processus non linéaires et qu'un nombre plus grand de VS est nécessaire pour leur décomposition. Le dernier volet de cette thèse étudie la possibilité d'utiliser le premier VS pour estimer les variations temporelles de la VI dans les simulations d'un MRC. Nos résultats montrent que le premier VS à lui seul ne permet pas d'anticiper des périodes caractérisées par une forte croissance de la VI dans les simulations du modèle.

Mots-clés: Variabilité interne, modèle régional du climat, vecteur singuliers, modèle linéaire tangent, instabilité hydrodynamique.

ABSTRACT

The internal variability (IV) is defined as the dispersion of model's solutions caused by its sensitivity to small differences in the initial conditions. Previous studies have shown that Regional Climate Models (RCM) IV fluctuates in time depending on synoptic events. This study focuses on the physical understanding of episodes with rapid growth of IV. An ensemble of 21 simulations, differing only in their initial conditions, was run over North America using version 5 of the Canadian RCM (CRCM). The IV is quantified in terms of energy of CRCM perturbations with respect to a reference simulation. The working hypothesis is that IV is arising through rapidly growing perturbations developed in dynamically unstable regions. If indeed IV is triggered by the growth of unstable perturbations, a large proportion of the CRCM perturbations must project onto the most unstable singular vectors (SVs). One set with ten SVs and ten sets with five SVs were computed for several 36-hour periods in the first twelve days of December 1992 with the goal to identify the orthogonal set of perturbations that provide the maximum growth with respect to the dry total-energy norm. CRCM perturbations were then projected onto the subspace of SVs and the IV tendency was analysed.

For the first five days of December, the IV tendency is dominated by the sources of instability. 36% to 85% of these sources are represented by the four or five most rapid SVs computed with the initial conditions from the CRCM reference state. Only on one instance was the projection dominated by the leading SV. In this particular cas, a large part of the IV growth is explained by initially small-amplitude unstable perturbations represented by the ten leading SVs, the SV subspace accounting for over 70% of the CRCM IV growth in 36 hours. The projection on the leading SV at final time is greater than the projection on the remaining SVs and there is a high similarity between the CRCM perturbations and the leading SV after 24 to 36 hours tangent-linear model integration. The vertical structure of perturbations revealed that the baroclinic conversion is the dominant process in IV growth for this particular episode. Although there appeared to be sources of instability in the domain, relatively small projections prevail in periods where the IV tendency is dominated by the IV export out of the analysis domain. The fact that the IV in such periods does not project into the directions of the most rapid SVs suggests that the IV growth is due to perturbations that are already in a growing state, and hence, characterised by a slower growth rate. As a consequence, a large set of SVs would be needed in order to decompose the CRCM sources of IV. For the setup used in this analysis, the leading SV alone does not appear to permit anticipating periods with larger than usual IV growth.

Keywords: Internal variability, regional climate model, singular vectors, tangent linear model, hydrodynamic instability.

INTRODUCTION

L'atmosphère est un système dynamique complexe caractérisé par une variabilité naturelle à différentes échelles spatiales et temporelles. Ces variations sont le résultat des interactions entre les différentes composantes du système climatique. Des fluctuations à basses fréquences sont induites dans l'atmosphère par des fluctuations similaires dans la température de la surface de la mer, l'humidité du sol ou les interactions avec la radiation solaire. Cependant, les oscillations à hautes fréquences caractéristiques aux phénomènes météorologiques sont le résultat des interactions non linéaires à l'intérieur de l'atmosphère. Les non linéarités représentent un ingrédient de base pour les systèmes chaotiques. Une propriété importante des systèmes chaotiques est leur sensibilité aux petites différences dans les conditions initiales (CI). Cette sensibilité a des conséquences profondes pour la prévisibilité du système.

Les modèles de climat sont des simulateurs numériques représentant des versions simplifiées de l'atmosphère. Ils visent à simuler l'état de l'atmosphère, soit sur tout le globe (les modèles atmosphériques de circulation générale - MCG) ou sur des régions limitées du globe (les modèles régionaux du climat - MRC). Il y a deux complications inévitables qui sont associées aux modèles de climat. Premièrement, les modèles représentent seulement des approximations de l'atmosphère réelle et donc ils ne sont pas parfaits. Deuxièmement, les modèles demandent des CI pour initialiser les simulations. Les variables utilisées pour simuler les principaux processus atmosphériques sont la température, le vent, la pression et l'humidité. Parce qu'il est impossible de connaître leurs valeurs exactes à chaque point de l'atmosphère, les CI du modèle contiennent toujours des erreurs. La nature chaotique du système atmosphérique fait en sorte que toute erreur ou perturbation dans le système non linéaire croîtra rapidement et par conséquent, après quelques jours, les oscillations simulées par le modèle seront différentes des oscillations réelles de l'atmosphère (Lorenz 1963, 1965). Donc, si plusieurs simulations d'un modèle sont lancées en utilisant des CI qui diffèrent par des quantités en accord avec le degré d'incertitude associée aux observations, elles vont donner des solutions différentes pour le même état atmosphérique qu'on vise à prévoir. En

pratique la prévisibilité des événements synoptiques est perdue après une dizaine de jours. On dit souvent dans la littérature que la mémoire des CI est perdue quand les solutions simulées par les membres d'un ensemble de simulations deviennent non corrélées l'une avec l'autre. La dispersion des solutions d'un ensemble de simulations causée par les différences dans les CI s'appelle variabilité interne (VI) du modèle.

En prévision numérique du temps, l'intérêt est de fournir aux modèles des CI les plus précises possibles. Ceci est réalisé à l'aide des différentes techniques d'assimilation de données qui fusionnent l'information apportée par un modèle numérique avec l'information apportée par les observations issues de stations de surface ou de satellites. De plus, des ensembles de simulations sont mis en place en utilisant différentes techniques spécialement conçues pour l'échantillonnage des erreurs potentielles dans les CI. La prévision d'ensemble est exploitée depuis le début des années 90 au centre américain de prévision météorologique NCEP (the National Centers for Environmental Prediction; Toth et Kalnay, 1997), au centre européen de prévision météorologique ECMWF (the European Centre for Medium-Range Weather Forecasts; Molteni *et al.*, 1996), et plus tard dans d'autres centres comme le Service Météorologique du Canada (SMC; Pellerin *et al.*, 2003). Les méthodes les plus populaires utilisées à l'heure actuelle dans l'initialisation des systèmes de prévisions d'ensemble sont : la méthode des vecteurs singuliers (VS; Molteni *et al.*, 1996; Palmer *et al.*, 2007;), la méthode du « cycle de culture » des perturbations (bred vectors en anglais; Toth et Kalnay, 1997), la méthode « filtre de Kalman d'ensemble » (Houtekamer et al., 2005) et la méthode « ensemble transform Kalman filter » (Wei *et al.*, 2006). Les deux dernières méthodes sont basées sur des méthodes d'assimilation de données.

Tel que précisé auparavant, la croissance des perturbations dans le temps va effacer la mémoire des CI. Par conséquent, à l'échelle climatique, la réponse du modèle à l'ajout d'une perturbation aux CI est indépendante de l'amplitude ou du type de perturbation. Cette propriété a été vérifiée et confirmée maintes fois. Par exemple Lucas-Picher *et al.* (2008a) ont initialisé les différents membres d'un ensemble MRC soit en décalant le moment d'initialisation, soit par l'utilisation de CI perturbées dans différents champs atmosphériques. Pour les modèles de climat, la manière dont les CI sont perturbées n'est pas importante. En général, pour les modèles de climat, l'ensemble est construit en décalant le moment initial de

chaque simulation par 6 ou 24 heures. Après une période d'ajustement (spin-up en anglais), la VI d'un MCG va atteindre le niveau de la variabilité naturelle climatique. Toutefois, la VI d'un MRC diffère de celle d'un MCG à cause des frontières latérales du modèle régional qui vont limiter à un certain niveau la croissance des perturbations et donc la VI.

Ces dernières années, plusieurs études ont été consacrées à la VI dans les MRCs (e.g. Giorgi et Bi, 2000; Rinke et Dethloff, 2000; Christensen *et al.*, 2001; Caya et Biner, 2004; Rinke *et al.*, 2004; Vanvyve *et al.*, 2007; Alexandru *et al.*, 2007; Lucas-Picher *et al.*, 2008a; Šeparović *et al.*, 2008; Rapaić *et al.*, 2010). Ces études montrent que la VI d'un MRC dépend de la taille et de la position du domaine, de la variable atmosphérique analysée, de la saison d'analyse et de la situation synoptique du jour. Par exemple, Lucas-Picher *et al.* (2008a) ont analysé la VI du MRC canadien pour un très grand domaine couvrant le continent nord-américain. Ils ont trouvé que la VI pour les champs de pression au niveau de la mer et de température en surface présente un cycle annuel avec un maximum au printemps. Cependant, la VI pour la précipitation a un très fort cycle annuel avec un maximum en été. Par ailleurs, l'étude de Rinke *et al.* (2004) effectuée sur un domaine circumpolaire arctique a montré que la VI pour la température présente un cycle annuel avec des maxima en automne et en hiver.

Une autre caractéristique importante de la VI des MRCs est sa dépendance de la grandeur du domaine d'intégration. Alexandru *et al.* (2007) ont comparé la VI d'un MRC pendant l'été 1993 pour cinq tailles différentes d'un domaine centré sur la côte est du continent nord-américain. Ils ont montré que la VI était plus faible pour un domaine plus petit, la contrainte imposée par les frontières latérales du modèle étant plus forte dans le cas d'un domaine plus petit. Ces résultats sont confirmés par Rinke et Dethloff (2000) et Rapaić *et al.* (2010) qui ont aussi analysé la VI pour différentes grandeurs de domaines situés dans l'Arctique et dans l'Atlantique du Nord, respectivement. Rapaić *et al.* (2010) ont calculé aussi séparément la VI pour les grandes échelles et pour les petites échelles en utilisant comme séparation une longueur d'onde de 1400 Km. Leur étude montre que la VI pour le champ de température est plus faible pour les grandes échelles que pour les petites échelles. De plus, la VI pour le champ de précipitations est plus forte que celle du champ de température.

La VI est aussi influencée par le temps qu'un élément d'air passe à l'intérieur du domaine. Le travail de Lucas-Picher *et al.* (2008b) met en évidence que la VI pour des champs qui sont pilotés aux frontières latérales du modèle (comme la pression au niveau de la mer, le vent et la température) est corrélée avec le temps de résidence, un long temps de résidence favorisant la croissance des perturbations. Toutefois, la corrélation est très petite pour la précipitation qui est un phénomène plutôt local et qui n'est pas piloté aux frontières latérales du modèle. Ainsi, la vitesse de l'écoulement à travers le domaine influence la divergence des solutions par l'intermédiaire du temps nécessaire aux perturbations d'arriver aux frontières latérales du modèle. Par conséquent, la position géographique du domaine a aussi une influence sur la VI. Une configuration avec un fort écoulement va pousser rapidement les perturbations vers les frontières latérales du domaine et leur croissance va être limitée. Cependant, une configuration avec une circulation en boucle à l'intérieur du domaine va favoriser la croissance des perturbations et donc la VI (Laprise *et al.*, 2008). Rinke *et al.* (2004) ont trouvé que la VI est beaucoup plus grande pour un domaine circumpolaire que pour des domaines de grandeurs semblables placés aux latitudes moyennes. La différence est due au fait que le domaine arctique est caractérisé par une faible circulation à travers les frontières latérales et un vortex quasi axisymétrique qui tient les perturbations captives à l'intérieur du domaine, spécialement en hiver et automne.

Nikiema et Laprise (2010) ont établi des équations de bilan de la VI d'un MRC pour deux variables atmosphériques : la température potentielle et le tourbillon relatif. Leur étude utilise un domaine situé sur la côte est du continent nord-américain et elle est consacrée à l'analyse de la VI pour l'été 1993. Les deux équations ont des termes similaires comme les termes du transport de la VI par l'écoulement de la moyenne d'ensemble et de covariance de fluctuations agissant sur le gradient de la moyenne d'ensemble. Leur analyse révèle que les épisodes de grande VI sont en général associés à un régime cyclonique. Les termes dominants responsables de la croissance de la VI pour la température potentielle sont les termes de covariance impliquant les fluctuations de température potentielle et le chauffage diabatique. Pour cette équation, l'énergie potentielle des fluctuations engendrée par les processus de condensation et convection est convertie en l'énergie cinétique des fluctuations. Pour le tourbillon relatif, les plus importants termes sont associés aux composantes horizontales de la moyenne d'ensemble, soulignant l'importance de l'écoulement horizontal.

Leurs résultats révèlent également que les épisodes de fortes diminutions de la VI se produisent lorsque les maxima de la VI sont proches de la frontière nord-est, la VI étant forcée par les conditions aux frontières latérales à diminuer jusqu'à zéro dans la région de pilotage.

On peut conclure que la divergence des solutions dans les simulations d'un MRC est le résultat de la compétition entre deux processus : la croissance des perturbations engendrées par la nature chaotique du système climatique et le forçage des conditions aux frontières latérales du modèle qui impose aux simulations la même solution. Ce forçage est fonction de la vitesse de l'écoulement à travers le domaine, qui dépend de la position et la grandeur du domaine ainsi que de la saison. De plus, la croissance des perturbations est modulée par les conditions synoptiques : des périodes d'activité synoptique importante vont favoriser une grande divergence des solutions et des périodes synoptiques calmes vont être caractérisées par une convergence des solutions.

Bien que plusieurs études aient été dédiées à l'analyse de la VI dans les simulations de MRCs, très peu ont été consacrées à l'analyse des processus physiques responsables de la croissance épisodique de la VI. Plusieurs hypothèses sont émises à l'heure actuelle pour expliquer cette croissance :

- Elle est stimulée par le déclenchement soudain des processus humides paramétrés, tel le relâchement de la chaleur latente lors de la précipitation stratiforme ou la convection;
- Elle est due au déclenchement d'une instabilité hydrodynamique, telle l'instabilité barocline dans les zones à forts gradients horizontaux de température;
- Elle est le résultat d'interactions non linéaires transférant la variance entre les échelles.

Cette thèse a comme objectif général d'étudier la croissance épisodique de la VI dans des simulations du Modèle Régional Canadien du Climat (MRCC). Plus précisément, la thèse s'intéresse à une période hivernale et a pour but de vérifier l'hypothèse selon laquelle la croissance épisodique notée dans la VI est due à des perturbations s'étant développées dans des régions atmosphériques dynamiquement instables et ayant un très grand taux de

croissance. Pour décrire l'espace instable des perturbations, nous avons fait appel à la technique des vecteurs singuliers (VS).

Les VS ont plusieurs applications dans les sciences atmosphériques. Ils sont souvent utilisés dans l'analyse des instabilités hydrodynamiques (Farrell, 1989; Borges et Hartmann, 1992; Farrell et Ioannou, 1996), dans le ciblage des observations (Buizza et Montani, 1999; Barkmeijer *et al.*, 2001; Wu *et al.*, 2009) et surtout en prévision numérique du temps pour générer les CI pour les systèmes de prévision d'ensemble (Molteni *et al.*, 1996; Palmer *et al.*, 2007; Yamaguchi et Komori, 2009). Les VS sont définis comme les perturbations qui croissent le plus vite dans un intervalle de temps fini et dans le contexte d'un modèle linéaire, et leurs amplitudes sont exprimées par des normes définies au temps initial et au temps final (Buizza, 1994).

Cette méthode de calcul de perturbations instables demande une version linéarisée du modèle (le modèle linéaire tangent, TLM), où certains processus physiques sont représentés de façon simplifiée. Des TLM avec différents degrés de complexité peuvent être utilisés pour calculer les VS, ainsi que différentes normes et intervalles d'optimisation, le choix étant en fonction du champ d'application.

Les VS constituent une base orthogonale de l'espace de phase du système et ils sont ordonnés conformément à leurs facteurs d'amplification : le premier VS est celui avec la plus forte amplification. Par conséquent, les directions associées aux premiers VS sont celles dans lesquelles une perturbation grandie le plus rapidement au sens de la norme utilisée dans leur calcul. Dans la plupart des applications et des études, les bases des vecteurs sont tronquées pour se limiter aux premiers 10 (ou 50 VS), qui représentent les perturbations les plus instables de l'écoulement. Ainsi, l'étude des VS apporte des informations sur la dynamique et la structure des instabilités avec de fortes croissances dans un temps fini, représentant en général une première étape dans l'analyse de l'évolution des perturbations dans l'atmosphère.

La présente recherche consiste donc à caractériser la croissance épisodique de la VI dans les simulations du MRCC à l'aide de la technique des VS. Dans ce contexte, les objectifs spécifiques suivants ont été définis :

1. Déterminer le choix optimal pour les paramètres requis par le calcul des VS sur le continent nord-américain en hiver.

II. Effectuer un ensemble de simulations avec la version 5 du modèle MRCC pour un mois d'hiver, afin de caractériser la VI et identifier un épisode avec une forte croissance.

III. Déterminer dans quelle mesure la croissance de la VI pour l'épisode sélectionné est expliquée par les perturbations représentées par les VS avec la plus forte amplification.

IV. Poursuivre l'analyse comparative de la VI avec les VS pour d'autres périodes avec différents niveaux de croissance dans la VI.

V. Évaluer si le plus rapide VS peut être utilisé à lui seul comme prédicteur de la VI.

Le travail effectué dans cette thèse est présenté en quatre chapitres dont trois correspondent à des articles écrits en anglais qui ont été soumis, ou le seront sous peu, pour publication dans différentes revues spécialisées. Cette introduction présente la problématique et le contexte général de la recherche, résume les objectifs et présente le plan de la recherche avec un bref aperçu des méthodes utilisées pour atteindre les objectifs.

Cette étude utilise comme méthode d'analyse la technique des VS. Le développement mathématique de la technique des VS est décrit en détail dans le chapitre I qui est constitué d'un article intitulé « *Singular Vectors in Atmospheric Sciences: A review* ». L'article va être soumis pour publication dans la revue *International Journal of Climatology* ou dans *Earth-Science Reviews*. Il est consacré à la synthèse d'une recherche bibliographique sur la technique des VS en sciences atmosphériques. L'objectif de cette recherche est de présenter les principales caractéristiques des VS et l'impact des différents choix de paramètres sur leurs structures et propriétés. Ainsi, l'accent est mis principalement sur le développement mathématique de la technique des VS et leurs caractéristiques. Cette synthèse a permis ensuite d'identifier les normes, l'intervalle d'optimisation et le degré de complexité du modèle linéaire tangent, les plus indiqués pour le calcul des VS dans le contexte de cette thèse (Objectif I).

Les objectifs II et III sont détaillés dans le deuxième chapitre qui est présenté sous forme d'un article intitulé « *Singular Vector Decomposition of the Internal Variability of the*

Canadian Regional Climate Model », qui a été soumis pour publication dans la revue *Climate Dynamics*. Dans ce chapitre, l'analyse a été concentrée sur un épisode marqué par une forte augmentation de la VI. Pour identifier les périodes de forte VI, un ensemble de 21 simulations qui diffèrent seulement dans leurs conditions initiales a été réalisé sur le continent nord-américain avec la version 5 du MRCC. Premièrement, l'ensemble de simulations MRCC est décomposé dans un état de base et une série de perturbations par rapport à cette référence. Ensuite, la VI est décrite en termes de l'énergie totale des perturbations. Pour extraire les perturbations les plus instables, une série de dix VS est calculée à partir de l'état de base des simulations du MRCC et les perturbations MRCC sont projetées dans les directions des VS. L'article offre une analyse comparative entre les caractéristiques de la VI dans les simulations du MRCC et celles des VS. Une attention particulière est accordée au premier VS. Nous avons aussi calculé le pourcentage dans lequel la VI du MRCC est expliquée par le premier VS et par la série des dix VS.

Dans la suite de cette étude, le troisième chapitre continue la recherche pour d'autres périodes de croissance dans la VI du MRCC, en utilisant encore dix séries de VS (Objectifs IV et V), chacune contenant les cinq premiers VS. En particulier, ce chapitre est consacré à l'analyse des périodes avec des croissances relativement petites par rapport au cas étudié dans le troisième chapitre. Le chapitre est constitué d'un article intitulé « *Can the leading Singular Vectors calculated from one simulation be used to estimate the variations of Internal Variability in an ensemble of Regional Climate Model simulations ?* » qui sera soumis pour publication dans la revue *International Journal of Climatology*. L'article a pour but de vérifier si les VS permettent d'anticiper les variations temporelles notées dans la VI de simulations du MRCC. Pour ce but, l'analyse est concentrée sur la variation totale de la VI et le domaine d'analyse est restreint à la partie centrale du domaine, loin des frontières latérales du modèle. Dans ce contexte, la fluctuation de la VI est le résultat de la compétition entre les processus qui génèrent la VI et le transport de la VI à l'extérieur du domaine d'analyse via le terme de la divergence du flux horizontal de VI par l'écoulement de reference. Une comparaison est réalisée entre la variation temporelle de la VI et la variation temporelle de la partie des perturbations MRCC projetées dans les directions des VS. Le rapport entre ces deux tendances indique dans quelle proportion la croissance de la VI est due à des instabilités représentées par les premiers VS. L'article met aussi en évidence la présence d'un transfert de

VI à l'extérieur du domaine d'analyse, ce qui est responsable de la décroissance de la VI à l'intérieur du domaine.

Le dernier chapitre résume les principaux résultats obtenus dans ces études et présente les perspectives de recherche ouvertes par ce travail.

CHAPITRE I

LES VECTEURS SINGULIERS DANS LES SCIENCES DE L'ATMOSPHÈRE : UNE SYNTHÈSE

Ce chapitre, rédigé en anglais, est présenté sous la forme d'un article qui sera soumis pour publication, soit dans la revue *International Journal of Climatology*, soit dans la revue *Earth-Science Reviews*.

Il présente une synthèse bibliographique sur la technique des VS en sciences atmosphériques. L'étude des VS apporte des informations concernant la dynamique et la structure des perturbations avec des croissances rapides dans un modèle linéaire et dans un temps fini. L'article présente le développement mathématique de la technique des VS et une synthèse de leurs applications dans les sciences de l'atmosphère. Une attention spéciale est accordée à la sensibilité des VS au choix des différents paramètres comme le temps d'optimisation, les normes, la résolution horizontale et le modèle linéaire tangent.

SINGULAR VECTORS IN ATMOSPHERIC SCIENCES: A REVIEW

by

Emilia Paula Diaconescu [*,1,3], Ayrton Zadra [1,2,3] and René Laprise [1,3]

[1] *Canadian Network for Regional Climate Modelling and Diagnostics*

[2] *Meteorological Research Division, Environment Canada, Montreal, Canada*

[3] *Centre ESCER (Étude et Simulation du Climat à l'Échelle Régionale), University of Quebec at Montreal, Montreal, Canada*

to be submitted to

International Journal of Climatology

or

Earth-Science Reviews

August 2011

[*] Corresponding author: Emilia Paula Diaconescu
Department of Earth and Atmospheric Sciences
Université du Québec à Montréal (UQAM)
P.O. Box 8888, Succ. Centre-ville
Montréal, Québec
CANADA H3C 3P8

Abstract

During the last decade, singular vectors (SVs) have received a lot of attention in the research and operational communities especially due to their use in ensemble forecasting and targeting of observations. SVs represent the orthogonal set of perturbations that, according to linear theory, will grow fastest over a finite time interval with respect to a specific metric. Hence, the study of SVs gives information about the dynamics and structure of rapidly growing and finite-time instabilities representing an important step toward a better understanding of perturbations evolution in the atmosphere. This paper reviews the SV formulation and gives a brief overview of their recent applications in atmospheric sciences. A particular attention is accorded to the SV sensitivity to different parameters such as optimization time interval, norm, horizontal resolution and tangent linear model, various choices leading to different initial structures and evolutions.

Keywords: singular vectors; growth rate; tangent linear model.

1.1 Introduction

This review article is aimed at scientists such as climatologists who, while not being experts in Numerical Weather Prediction (NWP), would like to gain some understanding in singular vector (SV) use. The number of studies implying the use of SVs has increased considerably over the last decade. To give a focus and to keep the article to moderate length, we concentrate in particular on the SV formulation and properties.

SVs can be defined in terms of the singular value decomposition of an operator (the so-called forward tangent linear operator) and can be physically interpreted as a set of fastest growing perturbations. The concept of SVs was first introduced by Lorenz (1965) in his analysis of forecast error growth in dynamical systems and developed later by Lacarra and Talagrand (1988). The optimization problem consists in finding the perturbations from a time-evolving model-generated basic state that have maximum growth (or amplification) in a finite interval of time. The solution can be reduced to a singular value decomposition of a suitable operator and the resulting SVs provide an orthogonal set of optimal perturbations. They represent the orthogonal set of perturbations that, according to linear theory, will have the maximum growth over a finite time interval with respect to a specific metric.

The growth of perturbations has been initially studied in simple models with idealised time-independent basic states assuming that the solution of the linear perturbation equations can be expressed as a superposition of orthogonal functions, i.e. normal modes with fixed structure and amplitude varying exponentially with time. However, Farrell (1982; 1985) while studying the growth of perturbations in quasi-geostophic models, found that it is also possible to identify perturbations with growth exceeding, for a limited time, those typical of the fastest growing normal mode. These perturbations have a three-dimensional structure that changes in time, and they were named "non modal perturbations" to distinguish them from the former ones, which have fixed structure during their time evolution and amplitude changing exponentially with time (Montani and Thorpe, 2002). Those studies used stationary basic states. SVs are « finite-time instabilities » (Molteni and Palmer, 1993) that are computed from a time-evolving basic state without assuming a normal-mode solution.

Initially, SV theory has been applied to a number of idealized studies within the framework of quasi-geostrophic models; it was used for example by Borges and Hartmann (1992) and Molteni and Palmer (1993) in barotropic and baroclinic atmosphere investigations. At the beginning of the 1990s, SV technique began to be computed for primitive equation models (e.g. the atmospheric general circulation models AGCMs). SVs have since been the subject of numerous studies. They have been used in atmospheric predictability studies to determine for example perturbations that are seed for mid-latitude synoptic systems (e.g., Palmer et al., 1994; Hakim, 2000; Descamps *et al.*, 2007) and in forecast and analysis error estimation (e.g., Ehrendorfer and Tribbia, 1997; Gelaro *et al.*, 1997). In the 2000s, SVs began to be used daily to construct the initial perturbations in Ensemble Predictions System (EPS) of several forecast centres as the European Centre for Medium-Range Weather Forecasts (ECMWF) and Japan Meteorological Agency. SVs capture the dynamically most unstable perturbations. Because they give the directions along which forecast error grows the fastest, they are considered to lead to optimal forecast spread and yield maximum information about the probability density function of the model state at a future time (Ehrendorfer and Tribbia, 1997). This property made them a very good candidate in producing the initial perturbations in EPS. Recently, they have been employed to perturb the initial conditions in coupled ocean–atmosphere models of El Niño–Southern Oscillation (ENSO) (Kleeman *et al.*, 2003). Another recent use of SVs is to detect "sensitive" parts of the atmosphere for targeting adaptive observations (e.g. Palmer *et al.*, 1998; Buizza and Montani, 1999; Langland, 2005; Buizza *et al.*, 2007b); by identifying the regions with a large sensitivity to small perturbations they point to where additional observations have the potential to significantly improve weather forecasts (Kim and Jung, 2009a).

The purpose of this paper is to provide an overview of the SV technique. We first review the mathematical formulation of SVs in Section 1.2. The linearization process around a time-evolving model-generated basic state leads to the tangent linear model (TLM). Linearized models have been developed first for the adiabatic part of the forecast model using a simplified scheme for the vertical diffusion. As a consequence, SVs have been obtained initially for dry models (e.g., Buizza *et al.*, 1993; Ehrendorfer and Errico, 1995; Buizza and Palmer, 1995). However, the process of error growth in the atmosphere depends not only on dry dynamics but also on moist diabatic processes such as condensation,

evaporation, and moist convection. The reasons for not including moist processes in earlier studies are due to the difficulties involved in finding suitable linear descriptions of moist processes (Errico and Raeder, 1999) especially because the parameterisations of subgrid-scale physical processes introduce "thresholds" that make the model discontinuous. The characteristics of "dry" and "moist" SVs are presented in Section 1.3, followed by a discussion regarding the SV dependence on the horizontal resolution and optimization time interval. Section 1.4 gives some examples of SV applications in atmospheric sciences. The paper concludes with a brief discussion of the main results that were presented in the review.

1.2 Mathematical development of the singular vectors

The mathematical development of SVs has been described in detail first by Lacarra and Talagrand (1988) and later by Buizza *et al.* (1993) and Kalnay (2002, ch 6.3). An equivalent mathematical development is presented also by Palmer *et al.* (1998) using index-based tensor formalism instead of the more conventional matrix notation. In this section, we review the theoretical bases and concepts using the matrix notation.

Consider a non-linear model (**H**) describing the atmospheric system. The variables needed to represent the atmospheric state of the model, such as temperature, wind and surface pressure, are collected as a column matrix called the state vector **X**. The time evolution of the state vector **X** (i.e. the model tendency equation) can be written in the symbolic form:

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X}),\ \mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix},\ \mathbf{F} = \begin{bmatrix} F_1 \\ \vdots \\ F_n \end{bmatrix}, \tag{1.1}$$

were $\mathbf{X} \in R^N$ denotes the N-dimensional state vector and $\mathbf{F}(\mathbf{X}) \in R^N$ its tendency which includes the dynamical and the physical parameterizations contributions.

Let $\mathbf{X}(t)$ be a solution of the Eq. 1.1. Finding the solution consists in generating a trajectory from an initial point $\mathbf{X}(t_0)$ to $\mathbf{X}(t)$ or to integrate Eq. 1.1 from t_0 to t. This is equivalent to looking at the model (the nonlinear propagator **H**) as a mapping of the initial-time vector $\mathbf{X}(t_0)$ onto a vector of predictions $\mathbf{X}(t)$:

$$\mathbf{H}: \mathrm{R}^n \rightarrow \mathrm{R}^n \qquad \mathbf{X}(t_0) \rightarrow \mathbf{X}(t) \tag{1.2}$$

$$\mathbf{X}(t) = \mathbf{H}(\mathbf{X}(t_0)). \tag{1.3}$$

The process of finding the optimal perturbations for a given basic state starts with the linearization of the non-linear model around the basic state, defined as the solution of the non-linear model or the trajectory in the space of states. Let *x(t)* be a small perturbation from the non-linear model trajectory **X**(t). Equation 1.1 can be written using the first-order Taylor-Young formula in the vicinity of the basic state **X**(t):

$$\frac{d}{dt}(\mathbf{X}(t)+x(t)) = \mathbf{F}(\mathbf{X}(t)+x(t)) = \mathbf{F}(\mathbf{X}(t)) + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right|_{\mathbf{X}(t)} x(t) + O(x^2(t)) \tag{1.4}$$

For short time internals and small perturbations, the terms $O(x^2(t))$ can be neglected, and, after the substraction of (1.1) from (1.4), the equation becomes

$$\frac{dx}{dt} = \left.\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right|_{\mathbf{X}(t)} x = \mathbf{A}_{\mathbf{F}} x \tag{1.5}$$

were $\mathbf{A}_{\mathbf{F}}$ is the Jacobian of **F**: $(\mathbf{A}_{\mathbf{F}})_{jk} = \dfrac{\partial \mathbf{F}_j}{\partial \mathrm{X}_k}$. Equation 1.5 defines the tangent linear model (TLM) equation. Its integration gives the evolved perturbation *x(t)* from any initial perturbation $x(t_0)$ via an integration of the tangent-linear model:

$$x(t) = \mathbf{L}(t_0, t)x(t_0) \tag{1.6}$$

Therefore, the operator $\mathbf{L}(t_0,t)$ that is named the forward tangent linear, or the linear propagator, can be regarded as a mapping of the initial time perturbation $x(t_0)$ onto the evolved perturbation *x(t)*. Equation (1.6) can be obtained equivalently by linearizing equation (1.3):

$$x(t) = \left.\frac{\partial \mathbf{H}}{\partial \mathbf{X}}\right|_{\mathbf{X}(t)} x(t_0) = \mathbf{L}(t_0, t)x(t_0) \tag{1.7}$$

Because $\mathbf{L}(t_0,t) = \left.\dfrac{\partial \mathbf{H}}{\partial \mathbf{X}}\right|_{\mathbf{X}(t)}$, it is also named the differential or first derivative of **H** (the

non-linear model) at point **X**(t). The equation shows that **L**(t_0,t) depends on the non-linear trajectory **X**(t), which evolves in time, but not on the perturbation ***x(t)***.

The second step in the optimal perturbation problem is to find those perturbations that have maximized amplitude growth over a finite interval of time. This can be done by finding the SVs of the forward tangent linear **L** through a singular value decomposition, which states that every matrix **L** can be decomposed as

$$L = W \Lambda Y^*, \tag{1.8}$$

where Λ is a diagonal matrix with non-negative real numbers on the diagonal. $\mathbf{W}=(w_1, w_2, \ldots, w_i, \ldots, w_n)$ and $\mathbf{Y}=(y_1, y_2, \ldots, y_i, \ldots, y_n)$ are orthonormal matrices and **Y*** denotes the conjugate transpose of **Y**. The columns of **Y** are named the right singular vectors (or the initial SVs) of **L**, while the columns of **W** are named the left singular vectors (or the final SVs) of **L**. The diagonal elements of Λ, diag($\lambda_1, \lambda_2, \ldots, \lambda_i, \ldots, \lambda_n$), are the singular values and they are ordered so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_i \geq \cdots \geq \lambda_n$. The relationship between each « *i* » initial and final SV is expressed by:

$$\mathbf{L}(t_0,t)\, y_i = \lambda_i w_i \tag{1.9}$$

Singular value decomposition and eigenvalue decomposition are closely related because the SVs of matrix **L** can be also obtained as the eigenvectors of **L*L**. Generally, the operator **L** is not normal ($\mathbf{L}^*\mathbf{L} \neq \mathbf{L}\mathbf{L}^*$). Normal operators (**S*S=SS***) are important because they can be diagonalized and spectrally decomposed in terms of orthogonal eigenvectors. However, because **L** is not generally normal, a traditional empirical orthogonal function (**EOF**) analysis for **L** yields eigenvectors that are not orthogonal to one another. Nevertheless, the operator **L*L** is symmetric and positive-definite, and therefore it is a normal operator with mutually orthogonal eigenvectors. The eigenvectors of **L*L** (**LL***) are the right (left) SVs of **L** and the eigenvalues of **L*L** (and **LL***) are the square of the singular values of **L**:

$$\mathbf{L}^*\mathbf{L}\, y_i = \lambda_i^2 y_i \tag{1.10}$$

$$\mathbf{L}\mathbf{L}^*\, w_i = \lambda_i^2 w_i \tag{1.11}$$

Consider now the problem of finding the perturbations with maximized amplitude

growth over a finite interval of time. In order to measure the magnitude of perturbations, we need to define a norm in the space of solutions. We shall refer to $\|\cdot\|_E$ as the **E**-norm and we shall assume that the solution space is an inner-product space so that the norm can be defined via an inner product,

$$\|x\|_E^2 = \langle x; Ex \rangle = x^T Ex, \tag{1.12}$$

where **E** is a matrix operator that defines the specific form of the inner product. It is important to note that the choice of metric is not unique; but **E** defines a valid norm if and only if **E** is Hermitian positive definite.

For any linear operator **L** on an inner-product space, there is a unique operator **L***, called the adjoint of **L** (conjugate transpose of **L**) with respect to the norm **E**, such that $\langle x; \mathbf{L}y \rangle = \langle \mathbf{L}^* x; y \rangle$ for all *x, y* in the space. For systems with real variables, the adjoint and the transpose are identical ($\mathbf{L}^*=\mathbf{L}^T$).

Different norms can be considered at the initial and final times. Consider now a perturbation with the norm at initial time $\mathbf{E}_0$

$$\|x(t_0)\|_{E_0}^2 = \langle x(t_0); E_0 x(t_0) \rangle \tag{1.13}$$

and at final time with the norm $\mathbf{E}_t$

$$\|x(t)\|_{E_t}^2 = \langle x(t); E_t x(t) \rangle \tag{1.14}$$

We are interested in the initial perturbations that maximize the ampification factor in the interval of time *(t_0, t)*, named optimization time interval (OTI). The amplification factor is sometimes named in literature as the growth rate, and it can be measured as the norm at final time divided by the norm at initial time:

$$\lambda^2 = \frac{\|x(t)\|_{E_t}^2}{\|x(t_0)\|_{E_0}^2} \tag{1.15}$$

Using Eq. 1.6 and the definition of adjoint operator, the amplification factor can be expressed as

$$\lambda^2 = \frac{\langle \mathbf{L}x(t_0); E_t \mathbf{L}x(t_0) \rangle}{\langle x(t_0); E_0 x(t_0) \rangle} = \frac{\langle \mathbf{L}^* E_t \mathbf{L}x(t_0); x(t_0) \rangle}{\langle x(t_0); E_0 x(t_0) \rangle} \tag{1.16}$$

Since E_0 and E_t are Hermitian positive definite, they are self-adjoint ($E_0^* = E_0$; $E_t^* = E_t$). The fastest growing perturbations are obtained by maximizing (1.16), and the solution is given by the following generalized eigenvalue problem:

$$(\mathbf{L}^* E_t \mathbf{L}) y_i(t_0) = \lambda_i^2 E_0 y_i(t_0), \tag{1.17}$$

If different norms are used at initial and final times ($E_0 \neq E_t$), the form equivalent to the Eq. 1.10 is obtained by using the variable transformation, $y_i(t_0) = E_0^{-\frac{1}{2}} \gamma_i(t_0)$:

$$(E_0^{-\frac{1}{2}} \mathbf{L}^* E_t \mathbf{L} E_0^{-\frac{1}{2}}) \gamma_i(t_0) = \lambda_i^2 \gamma_i(t_0) \tag{1.18}$$

In Eq. 1.10, $y_i(t_0)$ are the eigenvectors of $\mathbf{L}^*\mathbf{L}$ and the initial singular vectors of **L**. By analogy, the eigenvectors of $E_0^{-1/2}\mathbf{L}^*E_t\mathbf{L}E_0^{-1/2} = (E_0^{-1/2}\mathbf{L}E_t^{1/2})^*(E_t^{1/2}\mathbf{L}E_0^{-1/2}) = \mathbf{L}_s^* \mathbf{L}_s$, are the initial singular vectors of $\mathbf{L}_s = E_t^{1/2}\mathbf{L}E_0^{-1/2}$, hence, the perturbations with maximized amplification factor in the interval of time *(t_0, t)*. The amplification factors, λ_i^2, are given by the eigenvalues of the generalized eigenvalue problem or by the square of the singular values (λ_i).

The initial SVs form an $\mathbf{E_0}$-orthonormal basis and are ordered according to their amplification factors (λ_i^2) with the fastest growing structure being the first SV. The first SV is also known as the leading SV. Note that the adjoint of $\mathbf{L}(t_0,t)$ is $\mathbf{L}^*(t,t_0)$. The adjoint reverses the direction of time propagation. Therefore the left-hand side of the equation 1.17 involves one integration forward with the TLM, followed by one integration backwards with the adjoint model (ADM). Once the initial-time SVs are obtained, the corresponding final-time SVs (or evolved SVs) can be computed using $\gamma_i(t) = \mathbf{L}_s \gamma_i(t_0)$. The evolved SVs form an $\mathbf{E_t}$-orthogonal set at optimization time. Note also that SVs are computed following a time-evolving trajectory.

SVs can be computed using the same or different norms at initial and final time. Also,

limited-area norms can be used in order to obtain SVs over a geographically restricted region. Hartmann *et al.* (1995) have used a projection operator that sets the vector to have zero values on grid points outside the region of interest. If a projector operator is used at final time, the SVs are solutions of the following generalized eigenvalue problem:

$$(\mathbf{L}^{*}\mathbf{P}^{*}E_t\mathbf{P}\mathbf{L})y_i(t_0)=\lambda_i^2 E_0 y_i(t_0) \tag{1.19}$$

where $\mathbf{P}$ is the projection operator and $\mathbf{P}^*$ is its adjoint. Equation 1.19 can be written using the variable transformation, $y_i(t_0)=E_0^{-\frac{1}{2}}\gamma_i(t_0)$, in the form

$$(E_0^{-\frac{1}{2}}\mathbf{L}^{*}\mathbf{P}^{*}E_t\mathbf{P}\mathbf{L}E_0^{-\frac{1}{2}})\gamma_i(t_0)=\lambda_i^2\gamma_i(t_0)\,. \tag{1.20}$$

The use of the local projector operator $\mathbf{P}$ is relevant especially for the limited-area models and in the particular cases of targeting observations. These SVs are also referred to as "targeted SVs". For complex primitive-equation models, the matrix $\mathbf{L}$ exists only as an operator in the form of a computer code. In this case, the SVs are most easily obtained using an iterative Lanczos algorithm that does not require an explicit representation of the model operator (Buizza *et al.*, 1993; Errico, Ehrendorfer and Raeder, 2001).

Usually SVs are normalized to have unit norms at initial time. Hence, the evolved leading SV will give the direction in the phase space with maximum amplification or growth. This property constitutes the base of SV used in EPS. In NWP, the forecast ensemble must sample the space of potential errors in observations. SVs from the trajectory of a forecast model give the directions in which any small perturbation from the ICs will grow most rapidly. Therefore, the first n SVs permit to span the phase space in the n most unstable directions. These directions will define the n-dimensional unstable subspace of the system and they are considered to lead to sufficient forecast spread (Gelaro *et al.*, 1997).

The Eq. 1.17 shows that several choices must be made when SVs are computed:

- the norms at initial and final time,
- the optimization time interval (OTI),
- the trajectory,
- the TLM and ADM.

Because of the computational cost, usually the TLM and ADM are run at lower resolution and the trajectory comes from a low-resolution non-linear forecast. Several studies have analyzed the SV sensitivity to metric, OTI, TLM, trajectory and resolution (e. g. Buizza, 1998; Palmer *et al.*, 1998; Gilmour *et al.*, 2001; Buehner and Zadra, 2006). In the following, we present a short review of the SV dependence on the choice of norm.

As mentioned above the choice of metric is not unique. One restriction exists: it must be Hermitian positive definite. The most commonly used metrics in the literature are the enstrophy, total energy, kinetic energy and streamfunction variance norm. Palmer *et al.* (1998) have examined the dependence of SVs on these four norms for one case of middle-latitude SVs in December 1994. Here is their definition as presented in Palmer *et al.* (1998):

a) the enstrophy norm,

$$\frac{1}{2}\int \zeta_1\zeta_2 dVol \tag{1.21}$$

b) the streamfunction variance norm,

$$\frac{1}{2}\int (\Delta^{-1}\zeta_1)(\Delta^{-1}\zeta_2) dVol \tag{1.22}$$

c) the rotational kinetic energy

$$\frac{1}{2}\int (\nabla\Delta^{-1}\zeta_1)(\nabla\Delta^{-1}\zeta_2) dVol \tag{1.23}$$

d) the dry total energy norm

$$\frac{1}{2}\int \left\{(\nabla\Delta^{-1}\zeta_1)(\nabla\Delta^{-1}\zeta_2)+(\nabla\Delta^{-1}D_1)(\nabla\Delta^{-1}D_2)+RT_r\ln\pi_1\ln\pi_2+\frac{c_p}{T_r}T_1T_2\right\}dVol. \tag{1.24}$$

In these relations, ζ is the vorticity, D the divergence, π the surface pressure and T is the temperature. Indexes 1 and 2 refer to the vector 1 and vector 2 respectively, T_r is a reference temperature, R is the gas constant for dry air, and c_p is the specific heat for dry air at constant pressure. The integral denotes a finite sum over the entire model domain. The "dry" term in the name of the fourth norm highlights the fact that it does not consider moist perturbations. The dry total energy norm can be written also as the sum of three components:

kinetic energy, potential energy and surface-pressure energy. An energy norm that also accounts for the moisture disturbances is named moist energy norm; supplementary information about the moist energy norm can be found further in the cited articles.

Palmer *et al.* (1998) found that at initial time there is a great difference in scales between SVs computed with these four metrics:

- the streamfunction norm is characterized by lower to middle-troposphere small-scale baroclinic perturbations;
- the energy norm produces intermediate-scale baroclinic perturbations;
- the enstrophy norm has upper and lower levels large-scale perturbations.

The maxima of either enstrophy or streamfunction SVs do not coincide at initial time with those of the energy SVs. Although the initial-time SVs are quite sensitive to the choice of norm, the final-time SVs are much more similar, presenting a large-scale barotropic aspect. Same results are obtained by Frederiksen (2000) who has examined the dependence of SVs on the norm during periods of block development, using a two-level model with the upper level at 300 mb and the lower level at 700 mb.

Kalnay (2002, pp. 222-223) interprets the robustness of the SV to the choice of final-time norm as an obvious conclusion from the fact that all perturbations, including SVs, evolve towards the leading Lyapunov Vector that represents the direction in which maximum sustainable long-term growth can occur in a system without external forcing (therefore the system attractor). This idea is also sustained by Trevisan and Pancotti (1998) who have analysed the Lyapunov Vectors and Singular Vectors in the Lorenz system. Reynolds and Errico (1999) also showed that, into the context of a quasigeostrophic model, the SVs for optimization times of 5 days or longer converge tower a single pattern, the Lyapunov Vector. They pointed out that the different-metric final-time SV similarity grows very much for very long optimization times (10 to 40 days). The initial-time norm dependence of SVs is explained by Kalnay (2002) by the fact that SVs are initially outside the attractor, pointing to areas in the phase space where solutions do not naturally occur, and they rapidly rotate back into the attractor. The SV transient rapid growth in due to this rapid (one time only) rotation of the initial-time SVs toward the attractor (Szunyogh *et al.*, 1997).

Since the initial SVs based on a streamfunction-squared norm are too small scale, and the initial perturbations based on a vorticity-squared norm are too large scale, the most commonly used norm in EPS is the total energy norm. However, several studies point out that in EPS and adaptive observations, the most appropriate initial-time norm must be based on the inverse of the analysis-error covariance matrix, (e.g. Ehrendorfer and Tribbia, 1997; Palmer *et al.,* 1998). The argument is that the ICs in NWP are constrained both by the observing network and by the process of assimilating and analyzing the observations. Hence, the initial-time norm must account for spatial correlations of the initial errors and for the inhomogeneity of the observing network. Buehner and Zadra (2006) compared the analysis-error covariance and the dry total energy norms for extratropical SVs. They also found that the initial-time SVs shape is different from the total-energy SVs and that the shape of the evolved SVs is almost independent of the choice of norm. Barkmeijer *et al.* (1999) tested a norm given by the Hessian of the cost function of a variational assimilation scheme, which is constrained by an estimate of the analysis error covariance metric from 3D-Var. This norm is computationally five times more expensive than the total energy norm. Again, although the Hessian and total-energy SVs present different structures at initial time, there are no significant differences between the results obtained with an ensemble based on Hessian SVs and based on the total-energy SVs. Several authors (e.g., Palmer *et al.,* 1998; Leutbecher, 2007) concluded that among the simple metrics the total-energy metric is a reasonable first-order approximation to an analysis-error covariance metric, the final-time leading total-energy SVs approximating the major axes of the forecast-error covariance matrix (Buizza *et al.,* 1997). As a consequence, the most commonly used metric for both initial and final norm even in EPS, is the total-energy norm. However, the correct choice is indeed the analysis error covariance norm, which correctly evolves initial errors into final errors (Ehrendorfer and Tribbia, 1997).

The perturbations as physical structures have to satisfy particular dynamical properties such as "balance conditions", often used in forecast models to filter out the gravity waves. In their study of the balance of SVs, Montani and Thorpe (2002) pointed out that SVs calculated with enstrophy and kinetic-energy norms produce perturbations only with a vorticity component, the temperature part being forced to be zero. They remark that these kinds of perturbations are by construction unbalanced at initial time and therefore, they may produce

gravity waves in the first few hours of their time evolution. This means that a certain amount of energy might be radiated away instead of travelling with the perturbations. The fact that the total energy norm accounts for temperature, wind and surface-pressure disturbances qualifies it as a better candidate for the norm. However, Montani and Thorpe (2002) show that the linear balance equation is not satisfied for total-energy perturbations at initial time, but it is satisfied at optimization time. They point out that, in the context of a primitive-equation model, the SV evolution is such that the unbalanced part observed at initial time is progressively damped with time, so that at optimization time only the balanced part is present. This can also be interpreted as the result of the evolution of off-attractor initial-time SVs toward the on-atractor final-time SVs (Szunyogh *et al.*, 1997).

1.3 Structure and characteristics of SVs

As mentioned in the introduction, the tangent linear models have been developed first for the adiabatic part of a non-linear model because parameterizations of physical processes are highly nonlinear and sometimes discontonnous. Also, the inclusion of physical processes in TLM increases the computational cost of SVs considerably (Puri *et al.*, 2001). However, linearized models without physics produce unrealistic results. Buizza (1994) have underlined the necessity of the representation in TLM of vertical diffusion and surface drag (known as "dry" physics) in order to suppress shallow fast-growing structures near the surface, which are not of interest because they are strongly damped in the nonlinear integrations. Nowadays, the ECMWF TLM linearized physics accounts for vertical diffusion, gravity-wave drag, radiation, deep convection and large-scale condensation, the last two being known as "moist" physics. Zadra *et al.* (2004) analyzed the impact of dry and moist physics on SVs computed with the Canadian Global Environmental Multiscale (GEM) model. SVs computed using TLM that include moist physics are known as "moist" SVs, the opposite SVs being known as "dry" SVs. Two other important choices that must be made in computing SVs regard the TLM horizontal resolution and OTI. In the following, we present a brief description of the structure and characteristics of dry and moist SVs, followed by a review regarding the SVs sensitivity to the choice of TLM resolution and OTI.

1.3.1 Dry extratropical SVs

SVs structure depends on the norm chosen, various choices of the initial norm leading to different initial structures and evolutions. Because total energy represents the most commonly used norm, in the following we describe the structure of SVs computed with dry total-energy norm.

Usually the leading extratropical SVs have at initial time a westward tilt with height and a meridional phase tilt that diminish with time (e.g. Buizza, 1994; Montani and Thorpe, 2002; Coutinho *et al.,* 2004). The vertical westward tilt indicates baroclinically unstable perturbations that favour the conversion of available potential energy into kinetic energy, while the meridional tilt against the horizontal shear denotes a barotropic mechanism for SV growth characterized by the transformation of basic-state kinetic energy into the perturbation kinetic energy. Figure 1.1 presents an example of horizontal (Figures 1.1b and 1.1d) and vertical (Figures 1.1a and 1.1c) structure for a typical total-energy extratropical SV. This SV represents the leading SV from a set of ten SVs computed with the TLM of Canadian GEM model using an OTI of 36 hours and a final-time norm restricted to a region of the North American Continent. The basic state and the TLM had a horizontal resolution of 1° and 28 levels in the vertical. The ICs correspond to 5 December 1992 at 00:00 UTC. The initial- and final-time norms are the dry total energy norms (as defined in Zadra *et al.,* 2004) and the TLM includes the linearization of dynamical GEM core and a vertical diffusion scheme. Figures 1.1a and 1.1c show the temperature field at initial time, while the temperature field at final time can be visualized in Figures 1.1b and 1.1d. In 36 hours, the perturbation grew in amplitude and changed its shape: the vertical westward tilt changed into a slight eastward tilt and the horizontal tilt has disappeared. Usually, the same initial-time vertically tilted pattern characterizes the leading SVs in terms of wind or potential vorticity (PV) (see Montani and Thorpe, 2002 for the SV structure in terms of PV).

The total energy partitions into the kinetic, potential and surface pressure components for the first ten SVs at initial and final times are represented in Figure 1.2. The panels show that the ratio of kinetic to potential energy of SVs is approximately 1:3 at initial time and 5:1 at final time. The surface-pressure term represents only a small percentage of the total energy. Usually the total-energy middle-latitude SVs is characterized by a dominant potential

energy at initial time and a dominant kinetic energy component at final time (e.g., Gelaro *et al.*, 1998; Zadra *et al.*, 2004). The transformation from initial potential to final mainly kinetic energy can be interpreted as the rotation of the off-attractor initial-time SVs toward the on-attractor final-time SVs by adjustment processes. Because SVs were normalized to have unity total energy at initial time, the total energy at final time (Figure 1.2b) also illustrates the SV amplification factors (λ^2).

Other important information can be depicted from the total-energy vertical profile. Figure 1.3 displays the average total-energy vertical profile computed as the mean of the ten SVs total energy horizontally integrated over the globe. It can be seen that at initial time (dashed line), the energy maximum is located in the middle troposphere. Several studies have shown that dry total-energy SVs is usually located in the lower and middle troposphere at initial time (e.g. Buizza and Palmer, 1995; Montani and Thorpe, 2002; Coutinho *et al.*, 2004; Zadra *et al.*, 2004). During their growth, there is generally an upward energy transfer toward the jet level, sometimes accompanied also by a downward energy transfer toward the surface as can be seen in Figure 1.3 (solid line). The initial-time location is also a function of OTI: long OTIs yield SVs that initially have a lower location comparatively to shorter OTIs, which favour mid-troposphere structures. According to Coutinho *et al.* (2004), for a long OTI, the direct interaction with the surface tends to inhibit the initial growth and gives more distance for upward propagation and growth.

Montani and Thorpe (2002) used PV diagnostics to investigate the dynamical processes responsible for total-energy SV growth. They assumed that the perturbation PV can only be modified by exchanging potential vorticity with the basic state, which acts like a reservoir of PV, and they showed that SVs growth in terms of PV is due to the advection of the basic-state PV horizontal gradients. The growth in PV is usually greater at the tropopause where these gradients are more intense and the perturbation velocity components can interact effectively with the basic state and grow. By separating the perturbation in two parts, above and below the 500-hPa level, they found that the perturbations initially confined at low levels can propagate vertically more efficiently than those localized above 500 hPa, the interaction with the upper-level basic-state fields being much more effective in the former case. These results confirm the study of Badger and Hoskins (2001) who, in the framework of a

simplified Eady model, also identified the lower troposphere as the location where an initial perturbation can experience rapid growth. Montani and Thorpe (2002) found also that the final-time energy increases significantly at low levels by normal-mode-like growth through PV coupling, leading to near-surface perturbations. The energy growth at low levels has often the main final-time peak located just above the boundary layer as can be seen in Figure 1.3.

The energy spectra bring other information about the SV structure. Figure 1.4 (corresponding to Figures 3c and 3d from Zadra *et al.*, 2004) shows the initial-time and final time energy spectra, averaged among the first 45 SVs computed using the GEM TLM with a resolution of 3°, 28 levels in the vertical, and an OTI of 48 hours. The final-time dry total-energy norm was restricted to latitudes north of 30°N and the ICs were taken from the analysis of 16 February 2002. The energy spectra show an upscale energy transfer with a pronounced final-time spectral peak around wavenumbers 12-14, which is consistent with baroclinic disturbances at synoptic scales as noted in other studies (e.g. Buizza and Palmer, 1995; Gelaro *et al.*, 1998; Buehner and Zadra, 2006). This upscale energy transformation is one of the total-energy extratropical SV characteristics that distinguish them from the leading Lyapunov Vectors that have fixed structure during their time evolution (hence fixed spectra) and amplitude growing exponentially with time.

As previously mentioned, SVs structure depends on the norm. Barkmeijer *et al.* (1999) compared the 3D-Var Hessian SVs with total-energy SVs and found that the initial-time Hessian SVs differ considerably from the total-energy SVs. Figure 1.5 (corresponding to Figure 1 from Barkmeijer *et al.* (1999)) shows the total energy spectrum and the total-energy vertical distribution for the total-energy SVs and Hessian SVs. The initial-time Hessian SVs have much more power at large scales and are located especially at the upper levels as the final-time Hessian SVs. No upscale energy transfer is observed for Hessian SVs and their amplification is much slower. However, the authors remark that despite these differences, the leading 25 total-energy SVs and Hessian SVs explain nearly the same part of the two-day forecast error. This is due to the fact that both the Hessian SVs and the total-energy SVs are evolving toward the Lyapunov Vector. The fact that the Hessian SVs have the same shape at initial and final times and smaller growth rates, suggests that, unlike the initial-time total-

energy SVs, they are on the attractor at initial time, and thus represent real atmospheric perturbations.

1.3.2 Moist SVs

The importance of including moist processes in the computation of SVs was underlined by numerous papers (e.g., Ehrendorfer *et al.*, 1999; Coutinho *et al.*, 2004; Zadra *et al.*, 2004; Hoskins and Coutinho, 2005 for extratropical SVs; and Barkmeijer *et al.*, 2001; Kim and Jung, 2009a for tropical SVs). In this section we summarize the primary effects of moist physics on SVs.

Ehrendorfer *et al.* (1999) analyzed the effect of moist physics on extratropical total-energy SVs in the context of the National Center for Atmospheric Research (NCAR) Mesoscale Adjoint Modeling System (MAMS2). Two important properties were highlighted in their study. First, they show that a moist TLM leads to faster growth compared to the case in which only dry processes are considered. Secondly, there are new growing SVs that appear in the case of a moist TLM compared to a dry TLM. Hence, the moist processes not only modulate the dry SVs but they also add new mechanisms of error growth. The new SVs have maxima located in the lower troposphere and growth rates much larger than the growth rates of the dry SVs. These results highlight the necessity of a moist TLM in order to capture all structures that might potentially grow in a moist environment.

Another important concern is the choice of a norm that accounts for the moist part of SVs and its impact on SVs spectrum. Ehrendorfer *et al.* (1999) have used an extension of the widely employed dry total energy norm by adding a quadratic term based on the physical effect of condensation/evaporation on temperature:

$$E_m = \frac{1}{2}\frac{1}{A}\int_A \int_0^1 \varepsilon \left[\frac{L_c^{\,2}}{c_p T_r} q^2 \right] d\eta dA \,. \tag{1.25}$$

Here, q is the mixing ratio, T_r a reference temperature, c_p is the specific heat at constant pressure and L_h is the latent heat of condensation per unit mass. The constant ε is use to give different weights to the humidity term. The integral extends over the full horizontal domain A and vertical direction η. They found that the leading subspaces for dry

and moist total-energy norms are quite comparable and that the growth rates depend more on the choice of the basic state and linearized model (moist versus dry TLM) than on the choice of the norm (moist versus dry total-energy norm).

Similar results were obtained by Coutinho *et al.* (2004) and Hoskins and Coutinho (2005) using the dry and full physics TLM of ECMWF. Generally, the moist extratropical SV vertical structure is similar to the dry SV structure, both presenting a westward tilt with height. Coutinho *et al.* (2004) noted also that in general, the moist structures show less evidence of a horizontal tilt and therefore of barotropic conversion. The SVs computed with moist processes have a tighter and deeper structure and they peak at shorter wavelengths. This can be seen in Figure 1.6 (Figure 2 from Coutinho *et al.*, 2004), which presents the initial- and final-time total-energy spectra averaged among the first ten SVs for experiments with dry and moist TLMs at two horizontal triangular spectral truncations (T42 and T63) and two OTIs (24 and 48 hours). The shift to higher wavenumbers is more evident for the experiment with a higher resolution. Hence the use of a TLM with high horizontal resolution is necessary in order to reveal all small-scale structures that might grow rapidly in a moist model.

In their study, Kim and Jung (2009a) compared the dry and moist tropical SVs computed with the Pennsylvania State University - NCAR Mesoscale Model (MM5) TLM at 100-km horizontal resolution and with dry and moist total-energy norms. As in the case of extratropical SVs, moist physics increase the growth rate of tropical SVs and cause smaller horizontal structures located in the lower troposphere. However, the initial-time tropical SVs energy is dominated by the vorticity component and not by the potential component. Figure 1.7 (corresponding to Figure 11a, b, i and j from Kim and Jung, 2009a) plots the vertical distribution of energy for the leading SV at initial and final time, for the experiments with dry and moist total-energy norms and TLMs. When the moist total-energy norm with full weighting (ε=1) is used (figure 1.7 b), the initial-time kinetic and potential components are transformed in the dominant humidity component at final time. Similar results were obtained by Barkmeijer *et al.* (2001), who analyzed the tropical moist SVs using the global ECMWF TLM at T42 horizontal resolution. They mentioned also that in the case of SVs computed for tropical cyclone (TC) prediction, the choice of the target area is crucial in determining the

location and properties of SV because a large target area results in many SVs that are not located in the cyclone area. Puri *et al.* (2001) studied the ECMWF tropical SVs in terms of the spread in cyclone tracks and intensities; they noted also that larger target areas lead to reduced spread and therefore to a reduced potential value for the forecast. However, for small target areas, a much larger spread is obtained. Kim and Jung (2009a) have used a small target area located on the TC centre. They remarked that even for such a small target area, the moist SVs have relatively stronger sensitivity in the vicinity of a TC than the dry SVs, which usually have remote locations from the TC centre associated for example with a mid-latitude trough.

1.3.3 The TLM horizontal resolution and the OTI

Besides the norm, there are two other important parameters that must be chosen when computing SVs: the TLM horizontal resolution and the OTI. Several studies have focused on the analysis of the SV sensitivity to the choice of these two parameters (e.g. Buizza, 1994; Buizza, 1998; Komori and Kadowaki, 2010).

As mentioned in the previous section, Coutinho *et al.* (2004) used dry and moist TLMs with two different horizontal triangular spectral truncations (T42 and T63) for computing extratropical SVs. Their results showed that the use of a moist TLM requires a higher horizontal resolution in order to detect moist perturbations that have small-scale characteristics. Buizza (1998) also computed extratropical SVs using ECMWF dry TLM at three different horizontal spectral truncations (T21, T42 and T63). He compared the corresponding SV subspaces to the forecast error and found that T42 and T63 SVs capture better the forecast error scales in eight out of eleven cases. The fact that T63 SVs do not outperform T42 SVs could indicate that more physical processes must be included in the tangent model versions at higher resolution. Nowadays, the operational ECMWF EPS uses a T42 dry TLM for computing extratropical SVs (Palmer *et al.*, 2007).

In a recent paper, Komori and Kadowaki (2010) investigated the impact of the resolution on tropical SV structure for three horizontal triangular spectral truncations with linear Gaussian grid: TL63, TL95 and TL159. The SVs were computed for the analysis fields with the Japan Meteorological Agency (JMA) dry TLM and ADM using a dry total-energy

norm localized at final time over the region of a TC at an early recurvature stage. They found that the resolution has a great impact on the structure of dry SVs targeted on TCs. First it was found that the shift towards higher wavenumbers from initial to final time is much clearer in the case of a higher horizontal resolution. Secondly, different resolutions emphasized different phenomena in the structure of the leading SVs. Figure 1.8 (corresponding to figure 2 from Komori and Kadowaki, 2010) shows the total-energy horizontal distribution for the initial-time leading SV superimposed on the analyzed streamlines at 500 hPa. At TL63, the energy has a high peak located northwest from the TC centre and associated with the mid-latitude trough. A secondary peak is located in the low troposphere and associated with the TC surrounding flow. At TL159, the primary maximum of energy emphasizes the east side area of the cyclone (the TC surrounding flow) in the low troposphere, while the upper-level far-northwest structure is much more reduced. Therefore, the resolution of the TLM and ADM changes the balance among the contributing mechanisms. The authors concluded that a certain degree of resolution in the TLM and ADM is required to properly detect sensitive areas around the TC.

It can be concluded that a higher horizontal-resolution TLM is desirable especially in the case of tropical SVs computation. However, given the high cost of a high-resolution TLM, the operational ECMWF uses in ensemble prediction of TC only the first five moist SVs computed at T42, for up to six tropical optimization regions. The moist total-energy norm is limited between surface and 500 hPa, emphasizing on the surface maximum that is associated with TC (Puri *et al.*, 2001; Palmer *et al.*, 2007; Buizza, 2010).

Concerning the SVs sensitivities to different OTIs, Buizza (1994) has analyzed the impact of OTI on extratropical SVs computed with T21 dry TLM with initial amplitudes comparable with analysis error estimates, and he found similar SVs for OTIs varying between 24 and 72 hours. On the other hand, Komori and Kadowaki (2010) compared the TL159 dry tropical SVs targeted to TC for two different OTIs: 24h and 48h. They found that the 48h SVs have a larger influence from mid-latitude troughs than the 24h SVs, which are located mostly in the TC surrounding flow. This indicates that the large-scale mid-latitude structures need longer OTI to develop, while the SVs surrounding the TC have a more rapid growth and a shorter OTI will emphasize them.

The choice of OTI is related to the question of the validity of the tangent-linear approximation. The fact that SVs are computed in a linear approximation context puts an upper bound on OTI fixed by the approximation validity. This is usually tested by comparing the growth of SVs in the full nonlinear model and in the TLM. Buizza (1994) highlighted an upper limit for OTI of the order of 48 hours for the dry extratropical SVs at T21.

However, Tanguay *et al.* (1995) pointed out that the TLM validity depends on the resolution. Generaly a higher resolution implies a reduction of the TLM validity. Reynolds and Rosmond (2003) showed also that the TLM validity is a function of scale and norm. Their study indicated that the linear assumption for small scales is broken after twelve hours, while the large scale remain significantly linear up to two or three days. As mentioned before, the moist SVs have higher growth rate and present smaller-scale structures. In their conclusion, Reynolds and Rosmond (2003) pointed out that, despite the fact that a moist TLM is a more accurate representation of the full nonlinear model, its results can be less relevant for small scales because of nonlinearities that dominate the small scales of the full model. Errico and Raeder (1999) analyzed the accuracy of a moist TLM and adjoint model; they found that the tangent-linear assumption is quantitatively accurate for OTI of 24 hours in regions of significant dynamical forcing (therefore with large scale structures), but only qualitatively accurate in regions dominated by moist convection, which usually presents small-scale structures.

We note that most SVs studies use an OTI of 48 hours for dry SVs and an OTI of 24 hours for moist SVs.

1.4 Applications

SVs are employed for many types of applications in atmospheric science. Such applications include forecast and analysis error estimation, ensemble forecasting, targeted adaptive observations, predictability studies and growth arising from hydrodynamical instabilities. These applications are based principally on the fact that SVs form a complete basis of orthonormal structures ordered according to growth potential. Hence, as indicated in Ehrendorfer and Errico (1995), any arbitrary initial perturbation ($x(t_0)$) can be decomposed in

terms of this complete basis by projecting the perturbation in the directions of SVs:

$$x(t_0) = \sum_{i=1}^{N} \alpha_i \hat{y}_i(t_0) \tag{1.25}$$

$$\alpha_i = \left\langle x(t_0), E_0 \hat{y}_i(t_0) \right\rangle \tag{1.26}$$

Here, $\hat{y}_i(t_0)$ are the initial-time SVs with unity norm, α_i are the respective projection coefficients and E_0 is the norm at initial time. Therefore, the amplitude of the initial-time perturbation can be written as

$$\left|x(t_0)\right| = \sqrt{\left\|x(t_0)\right\|^2} = \sqrt{\sum_{i=1}^{N} \alpha_i^2} \tag{1.27}$$

while the amplitude at final time will be

$$\left|x(t)\right| = \sqrt{\left\|x(t)\right\|^2} = \sqrt{\sum_{i=1}^{N} \alpha_i^2 \lambda_i^2} \tag{1.28}$$

The final-time magnitude of the perturbation depends on the projection coefficients α_i^2 weighted by the individual amplification factors λ_i^2. The complete basis of SVs contains growing, neutral and decaying SVs. In their study on the spectra of singular values in a regional model, Errico *et al.* (2001) used several sets with a very large number of SVs; they showed that a large part of SVs are slowly growing or decaying SVs, and that only a small fraction of the SVs are very fast growing structures. As a consequence, for the arbitrary perturbation expressed in equations 1.27 and 1.28, growth is expected only if the magnitudes of the projections on the growing SVs ($\lambda_i^2>1$) are sufficiently large with respect to those for decaying SVs ($\lambda_i^2<1$). Errico *et al.* (2001) pointed out that even if projection in the leading SVs is much greater than the projection in other individual SVs, the overall projection onto the slow growing and decaying SVs could overwhelm the projection onto the leading SVs because the number of slow and decaying SVs is very large. As a consequence, knowledge of the full spectrum of SVs (or the most part of it) should be of interest when studying

atmospheric perturbations.

Due to computational cost constraints, most studies with complex models use subsets of SVs limited to the first 5 to 50 most rapid SVs, out of a possible set of 10^6-10^7 or more. In the particular case of a growing perturbation, which projects uniformly on a small set of initial-time SVs, if only a few singular values are large, it is expected that at final time the projection onto the leading SVs will dominate the projection onto the other SVs (Hartman *et al.*, 1995). In their study on SV sensibility to the horizontal resolution, Buizza (1998) considered a set with only the first ten fastest growing SVs. The 10-dimensional subspace of the system's phase space at time t spanned by these SVs is:

$$G(t) = \left\{ \hat{y}_j(t), j = 1,10 \right\} \tag{1.29}$$

Only the fast growing part ($\tilde{x}(t)$) of an arbitrary perturbation ($x(t)$) will project into this subspace:

$$\tilde{x}(t) = \sum_{j=1}^{10} \alpha_j \hat{y}_j(t) \tag{1.30}$$

The ratio between the norms of the projected part and the full perturbation is usually used to measure the part explained by the SVs:

$$p(t) = \frac{\left\| \tilde{x}(t) \right\|^2}{\left\| x(t) \right\|^2} \tag{1.31}$$

This ratio is called the projection index. Buizza (1998) used this index to find the proportion in which the final-time leading ten SVs, computed with ECMWF TLM at T21, T42 and T63 horizontal resolution, can explain the two-day forecast errors for the period of 3 to 13 September 1993. The projection indices were rather small, with values between 20% and 45% and an average of 28%, suggesting that more than ten SVs are needed to describe the forecast error. A larger percentage was obtained by Barkmeijer *et al.* (1999) who projected the operational two-day forecast error onto 25 final-time Hessian SVs and 25 final-time total-energy SVs. Both types of SVs described nearly the same fraction in terms of total energy, which varies for the ten cases analyzed between 30% and 60%. However, there were other cases when a set of 25 total-energy SVs explained a very small percentage of two-day

Northern Hemisphere forecast error. For winter cases, Barkmeijer *et al.* (2003) obtained percentages varying between 0.1 and 0.22, while for summer cases the percentages had values between 0.08 and 0.19.

In their study, Gelaro *et al.* (1997) compared a set of 30 SVs at initial time to the analysis error field and to the forecast error sensitivity pattern obtained from an adjoint model integration. The forecast sensitivity fields indicate regions where analysis error leads to significant forecast error, rather than regions with large analysis errors. For example, if the analysis error was large in an area of weak dynamical instability, it may have a relatively unimportant contribution to forecast error (Molteni *et al.*, 1996). They showed that the analysis error projects with equal probability onto all initial-time SVs, and that 30 evolved dry-total-energy SVs at T42 resolution captured a large fraction of the covariance of the Northern Hemisphere sensitivity pattern in most of the cases examined. Gelaro *et al.* (1997) have grouped the analyzed cases into two categories: cases when the forecast had a low skill and cases when the forecast had an average skill. They pointed out that fewer SVs were needed in the low-skill cases than in average-skill cases. It is important to note that in the average-skill cases, the patterns were less localized and the projection was made onto the slowly growing final-time SVs and not on the leading SV, while the projection for the low-skill cases is dominated by the leading final-time SVs. The fact that a relatively small set of SVs captured a large part of the variance of forecast sensitivity fields qualifies it to be used to sample the ICs in ensemble forecasts. Until 2010, the ECMWF EPS uses a combination of 50 SVs, half initial SVs and half evolved SVs, for each of the extra-tropic hemispheres to compute the perturbations that are added and subtracted from the operational analysis.

There are two primary error sources in forecasts: errors in the estimation of the initial model state and errors due to imperfections in the model formulation. Ensemble forecasts with perturbed ICs deal with the initial state errors. Theoretically, the statistical properties of initial state errors can be described in terms of a high-dimensional probability density function (pdf) of the model state (Houtekamer, 1995). A forecast ensemble is based on the idea of finding a method to adequately sample the initial state pdf in order to obtain an ensemble of different but equally likely initial states (with the same mean and variance, in the ideal case). Furthermore, the resulting initial states are evolved using the model and result

into an ensemble of forecasts. In order to fully capture the dynamics of the atmospheric circulation, it is important that the sample of initial states results in a sufficient spread for the ensemble (Barkmeijer, 1996). In other words, in a good ensemble, the verification is a plausible ensemble member (Toth and Kalnay, 1993).

At ECMWF, the forecast ensemble for the first ten days have 51 members run at TL399 (Buizza *et al.,* 2007a). For the unperturbed member, the ICs are given by analysis error estimate provided by the ECMWF 4D-Var system. The 50 perturbed members have ICs modified by adding and subtracting 25 perturbations based on SVs computed from a six-hour forecast initialized six hours in advance (Buizza *et al.,* 2008). By using pairs of opposing perturbations, the mean (the first moment of the pdf) of the ensemble is initially identical to the non-perturbed member (the control forecast). SVs give the directions of the most rapid perturbation growth for a given measure. So, by sampling the phase space in the most unstable directions, the ensemble is supposed to account for a maximum fraction of pdf variance (the second moment of the pdf).

In the ECMWF EPS, the perturbations are defined over the two extra-tropical hemispheres and over up to six tropical regions with active TCs. 50 total-energy SVs are computed with the dry T42 TLM and an OTI of 48 hours for the northern extra-tropical hemisphere and another 50 SVs for the southern extra-tropical hemisphere. In order to generate some spread in the TC evolution, tropical SVs are computed with a moist TLM and a norm confined to the levels below 500 hPa and targeted to the few main regions where tropical storms are developing. Once computed, SVs are combined and re-scaled to create perturbations that cover most of the targeted regions without overlap, and to have amplitude comparable with analysis error estimate from the high-resolution data-assimilation system. A weakness of the ECMWF EPS consists in the sampling of the tropics (Buizza et al., 2008). The tropical EPS perturbations are computed only for a few selected regions and are based on only five initial-time SVs for each of these regions. Before 2010, the 48-hour linearly final-time SVs computed from a forecast started 54 hours in advance were also used in order to simulate large-scale structures that are present in the analysis as a result of errors that have already grown during the current and past data assimilation cycles. After 2010, a new set of perturbations generated from the Ensemble Data Assimilation (EDA) system replaced the

final-time SVs. Buizza *et al.*, (2008) shows that the use of this new set of perturbations increases the ensemble score especially over the tropics. The projection of the forecast error onto the orthonormal base defined using ten perturbed forecasts with EDA and initial-time SVs perturbations has a value of 50% for the 500 hPa geopotential height over the Northern Hemisphere extra-tropics at forecast time of 120 hours and 30% for the 850 hPa temperature over the tropics at the same forecast time. This result was obtained as an average over twenty cases spanning a 40-day period. A detailed description of ECMWF EPS can be found in Molteni *et al.* (1996) and information about later developments are presented in Buizza *et al.* (2007a), Palmer *et al.* (2007) and Buizza *et al.* (2008).

The Japan Meteorological Agency (JMA) runs ensemble prediction systems derived from the TL319 global model that also use SVs to perturb the ICs provided by the 4D-Var analysis:

- The One-Week Ensemble Prediction System (WEPS) once a day with 51 members;
- The Typhoon Ensemble Prediction System (TEPS) four times a day with 11 members.

The main difference between WEPS and ECMWF EPS consists in the targeted areas. In WEPS, moist SVs are targeted for the whole tropics and the dry SVs are targeted only for the northern hemisphere.

The TEPS became operational in February 2008 and has an aim to improve track forecast targeting for TCs in the western North Pacific and the South China Sea. For one forecast event, up to forty total-energy SVs are computed using TLM at T63 resolution and OTI of 24 hours:

- Ten dry SVs targeted for the mid-latitude;
- Up to thirty moist SVs targeted for up to three TCs at a time.

The moist SVs are targeted to a rectangle of ten degrees in latitude and twenty degrees in longitude with its centre at the forecast TC's central position. A detailed description of the JMA TEPS is given in Yamaguchi and Komori (2009).

Other forecast centres that use SVs to initialize ensemble forecasts are Météo-France and Bureau of Meteorology (BoM) in Australia.

SVs have been also employed to quantify the growth of perturbations in the atmosphere and oceans in order to explain the geophysical fluid dynamics of a particular climatic phenomenon such as cyclogenesis or ENSO.

Cheng *et al.* (2010) used SVs to study ENSO predictability for a period of 148 years. They have computed SVs for the Zebiak–Cane model of the tropical Pacific coupled ocean–atmosphere system with an OTI of six months for the sea surface temperature (SST) field; they found that the SST of the coupled leading SV has a west-east dipole structure oriented across the tropical Pacific, with one centre of action located in the east Pacific and the other in the centre Pacific. Their analysis was limited to only the first SV and small correlation has been found between the leading SV and ENSO predictability.

Descamps *et al.* (2007) questioned the ability of SVs to explain the cases of mid-latitude cyclogenesis. By subtracting a cyclone and its precursors from a basic state and comparing it with a set of ten SVs computed using dry total energy norms and OTI of 24 hours corresponding to the incipient stage of cyclogenesis, they found that the structure of a single SV has little to share with that of a real cyclone.

Other studies used SVs not to explain the cyclogenesis but to identify sensitive regions that can influence the evolution of a TC. A better understanding of these processes could help in future improvements of TC forecasts. An example is the study of Chen *et al.* (2009) who used SVs to obtain information about the dynamical processes that have an important impact on TC evolution. The composites for 72 cases with TCs show that the maximum initial-time SV is located at around 500 kms from the centre. Other initial-time SVs are situated in confluence regions generated by several systems such as the mid-latitude jet, the subtropical high and the TC. Therefore, it highlights these regions as areas with a delicate balance that influences the evolution of several synoptic systems. Kim and Jung (2009b) and Reynolds, Peng and Chen (2009) also used SVs to examine TC sensitivity during recurvature and the subsequent downstream impacts resulting from the interaction of the TC with the mid-latitude environment. They showed that the SVs situated in mid-latitude upper trough region become dominant as the TC recurves, while the SVs close to the TC centre play an important

role when the TC is far from recurvature. Kim and Jung (2009b) pointed out that the sensitive area around the TC centre is associated with warming in the mid-troposphere, while the sensitivity area under the upper trough are associated with strong baroclinicity and frontogenesis. Most of the initial-time remote SVs evolve to mid-to-lower troposphere structures that are co-located with the TC at final time.

SVs are also employed in targeted (adaptive) observation. Target observation strategies aim at identifying regions where additional observations have the potential to significantly improve weather forecasts. Changes to the ICs in these ''sensitive'' regions are expected to have a larger effect on the forecast skill than changes in other regions (Kim and Jung, 2009a).

Targeted observations are concerned especially with the extra-tropical and tropical cyclones because a large part of forecast errors are associated with them. The first targeted experiment was the Fronts and Atlantic Storm Track Experiment (FASTEX), which focused on the cyclones that form over the Atlantic Ocean and that are moving towards Europe. Other experiments targeted extra-tropical and tropical cyclones over the Pacific (e.g. NORPEX, the North Pacific Experiment; WSRP, Winter Storm Reconnaissance Program; THORPEX PARC, The multinational Observing System Research and Predictability Experiment for Pacific Asian Regional Campaign; DOTSTAR, Dropwindsonde Observation for Typhoon Surveillance near the Taiwan Region).

Several strategies have recently been proposed to evaluate the impact of targeting observations on the quality of forecasts (Rabier *et al.*, 2007). A detailed comparison of the different targeting techniques for the western North Pacific basin was given in Wu *et al.* (2009) and for the Atlantic in Majumdar *et al.* (2006) and Reynolds *et al.* (2007). One of the popular strategies is to use SVs to identify the ''sensitive'' regions. In this case, the norm of final-time SVs is confined to the geographical area (verification area) where the forecast must be improved and the initial time global SVs will indicate the "target area" where extra observations should be taken in order to reduce the forecast error inside the verification area (Buizza and Montani, 1999). Buizza and Montani (1999) suggested to diagnose the "target area" using a localization function based on vertically integrated total energy of SVs ($\boldsymbol{E}_j$) and defined as weighted average of the leading N SVs:

$$f(x)=\frac{1}{N}\sum_{j=1}^{N}\frac{\lambda_j}{\lambda_1}E_j(x)\,. \tag{1.32}$$

For each case, the target area is defined by the grid points x with the largest $f(x)$. Maps of forecast error variance reduction due to the use of an additional sounding at a given location agree with maps resulting from the former localization function. Buizza and Montani (1999) applied this function to identify "target areas" over the Atlantic in order to reduce the forecast error over Europe; they found that in some cases errors could be reduced by up to 13%.

In more recent papers, Buizza *et al.* (2007b), Cardinali *et al.* (2007) and Kelly *et al.* (2007) investigated the impact of removing targeted observations in the Pacific/Atlantic oceans on the two-day forecast error verified over North America/Europe. They found that observations taken in SV-target areas are more valuable than in randomly selected areas and that the value of targeted observations depends on the region, the season and the baseline observing system (data-rich or data-poor). For example, in a data-poor case, SV-targeted observations over the Pacific reduced the two-day forecasts error of 500-hPa geopotential height in the verification region by 27.5%, while for a data-rich case a reduction of 4.0% was obtained.

For TCs, several studies showed that despite the increased use of satellite data in the analysis of NWP models, additional dropsonde measurements of key variables such as wind, temperature, and humidity below cloud where satellite observing capabilities, are more limited can lead to improvements in track forecasts of the order of 10%–20% (Yamaguchi et al., 2009). Harnisch and Weissman (2010) investigated the benefit of TORPEX-PARC dropsonde observations in different locations with the ECMWF Integrated Forecasting System (IFS) experiments. As mentioned previously, two regions are often indicated by total-energy SVs as sensitive for a TC: an area in the vicinity of the storm and a remote area associated with a mid-latitude upper trough. Harnisch and Weissman (2010) found that the largest TC track forecast improvements are found for observations in the vicinity of the storm as indicated by SVs, but only a relatively small influence with a slight positive tendency is observed for dropsondes in remote regions.

Nowadays, in NWP a wide range of satellite observations are used to constrain the analysis (about 95% of the data employed in the 4D-Var system at ECMWF originate from satellites) and this volume of data is increasing. The huge stream of satellite data can be kept to affordable levels if satellite data density is increased only in a selective way (SV-based thinning). In this context, an important issue is to assess whether SVs can be used to identify areas where extra satellite data can be used to reduce analysis uncertainty and the forecast error (Langland, 2005)

1.5 Concluding remarks

This article is an introduction to and a review of the singular vectors formulation and properties. Singular vectors (SV) represent the orthogonal set of perturbations that, according to linear theory, will grow fastest over a finite-time interval with respect to a specific metric. Therefore, the study of SVs gives information about the structure and dynamics of rapidly growing perturbations and finite-time instability. SVs are computed for many types of applications in atmospheric science, such as forecast and analysis error estimation, ensemble forecasting, target adaptive observations, predictability studies and growth arising from instabilities.

Several choices must be made when SVs are computed: the norms at initial and final time, the optimization time interval (OTI), the trajectory, tangent linear model (TLM) and adjoint model (ADM). The most commonly used norm in various applications is the total energy norm and the OTI varies usually between 24 hours and 48 hours. The choice of OTI is conditioned by the validity of the tangent-linear approximation, which is function of scale. The small scales are dominated by nonlinear perturbations growth after 12 hours, while the large scales remain fairly linear out to 48 hours. The resolution of the tangent and adjoint models and OTI change the balance among the contributing mechanism in SV development. Therefore, the choice of parameters must be related to the field of application. For example, an OTI of 48 hours and a dry TLM at a lower horizontal resolution are usually employed for extratropical baroclinic perturbations, while moist perturbations need a moist TLM with higher resolution and an OTI of 24 hours in order to capture the small-scale characteristic of moist processes.

Another important issue in SV computation concerns the number of SVs necessary to capture the growing part of an arbitrary perturbation. While very unstable perturbations evolve usually into the direction of the most unstable SVs, a large number of SVs is needed to describe an arbitrary perturbation. Errico et al. (2001) indicated that the knowledge of most of the spectrum of SVs should be of interest but the computation of a very large number of SVs is not practical because it requires high computational resources.

Nowadays, the attention is focused on the TLM development to include more physical processes in order to approach the TLM to the non-linear model. The linearization of physical processes reveals a lot of difficulties due to the treatment of conditional formulations.

The proper choise of the norm is still an open question, because the key properties of initial-time SVs are strongly norm-dependent. For practical reasons, most studies continue to use the total-energy norm. However, the total-energy initial-time SVs are outside the atractor and, as a consequence, the projection of atmospheric perturbations (or forcast errors) on initial-time SVs is negligible. More efforts are required towards finding an initial-time norm that reveals initial-time SVs similar to atmospheric perturbations.

An alternative method in the study of stability properties of evolving flows is given by bred vectors (Kalnay, 2000). Bred vectors, as SVs, evolve also toward the Lyapunov vectors. They also represent finite-amplitude, finite-time vectors, but they do not require linearizing the model and are independent of the norm used to define the size of the perturbation. The breeding method consists of generating a control run of a nonlinear model for a short period of time (e.g., 6 h), then perturbing the atmospheric initial conditions and running the same model again for the same period of time (perturbed run). The difference between the two model runs at the final time is adjusted to the amplitude of the initial perturbation and is added to the new control initial condition. The method is then repeated. After a few days of iteration, the difference between the control and perturbed model runs represents a sample of fast-growing nonlinear perturbations (e.g., Szunyogh *et al.*, 1997; Kalnay, 2002).

Acknowledgments

This manuscript is based upon work supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Graduate Scholarship (CGSD3), as well as by the CRCMD Network supported by the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS) and the Ouranos Consortium. We thank Prof. Eugenia Kalnay for her comments on a previous version of the paper and Prof. Pierre Gauthier for pointing to us the references of Tanguay et al. and Rabier et al.

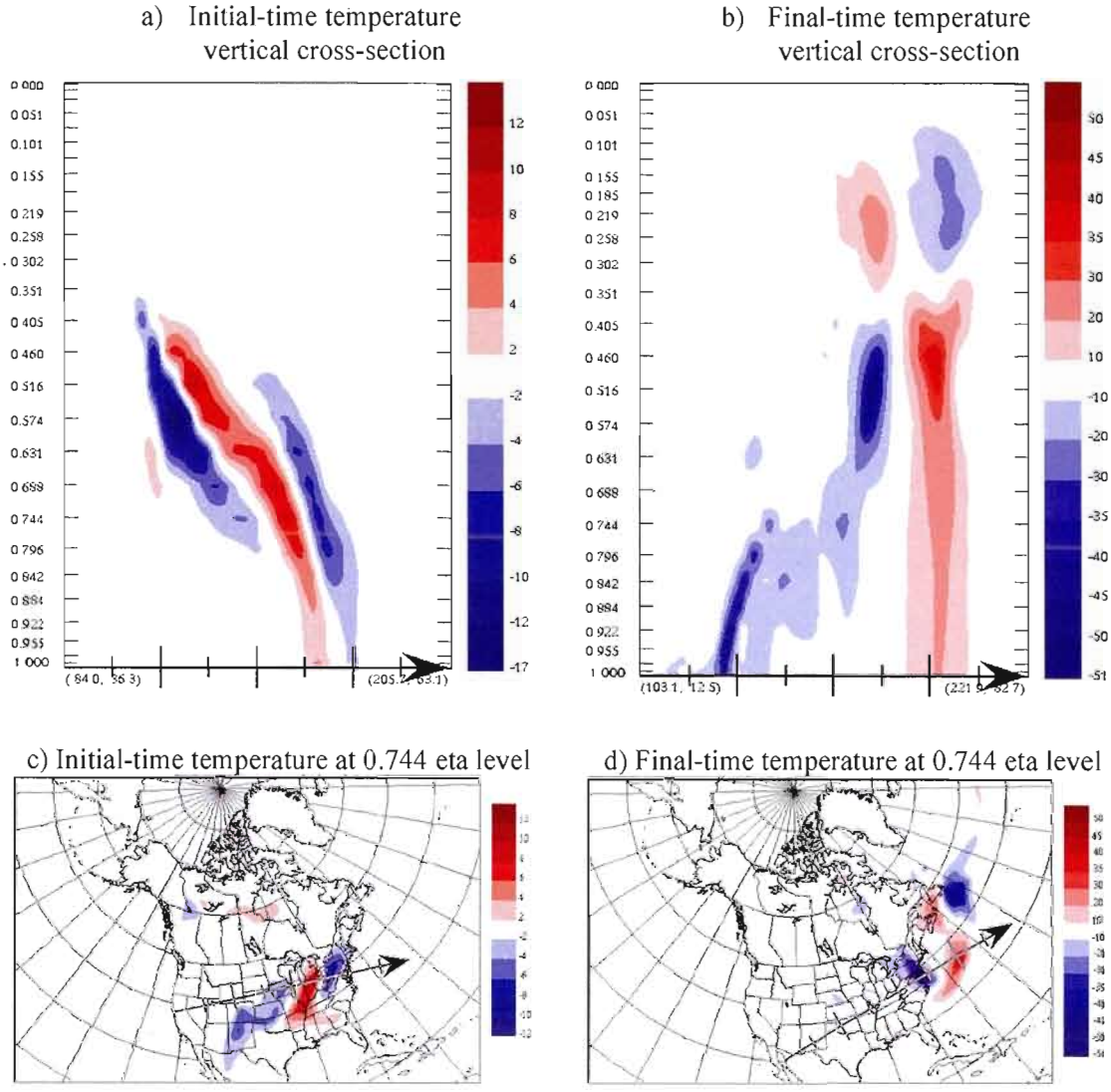


Figure 1.1: (a and b) Vertical cross-section and (c and d) horizontal cross-section at the 0.744 eta level of the temperature field corresponding to the first singular vector at (a and c) initial time and (b and d) final time. Note the different contour intervals between the (a,c) and (b,d) panels.

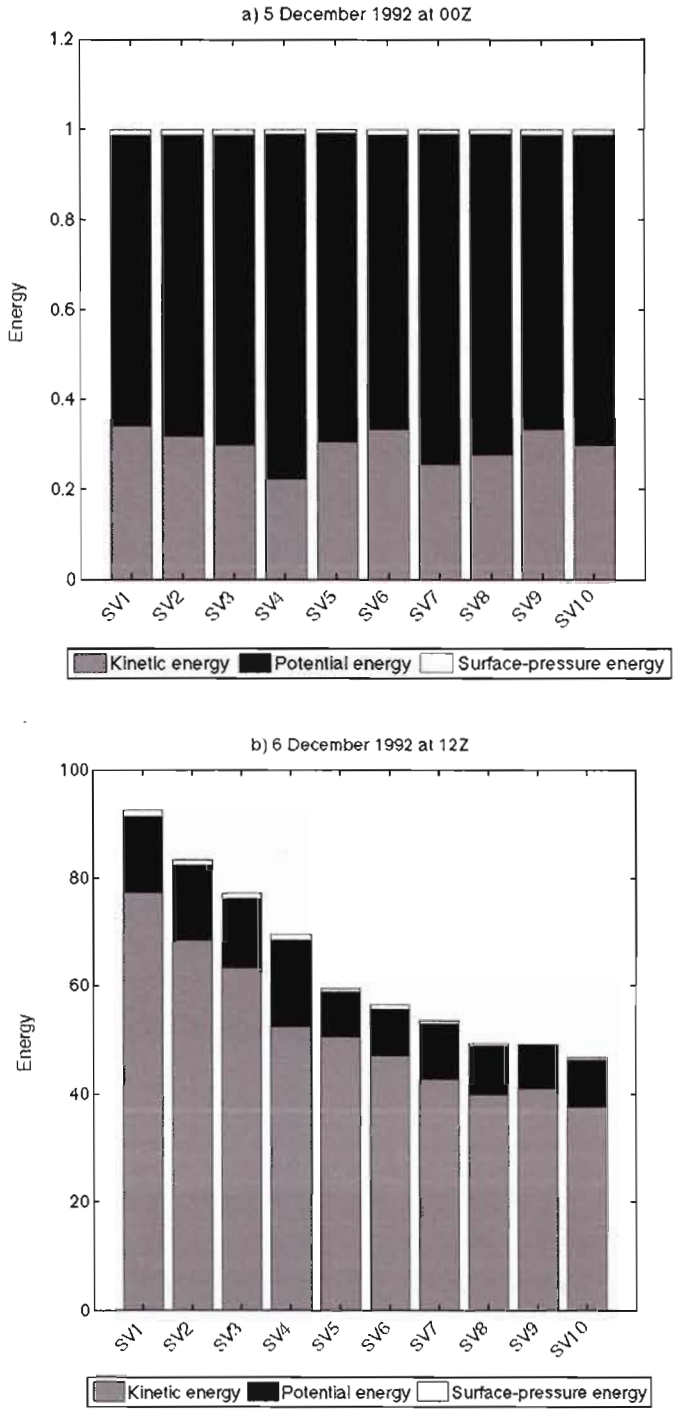


Figure 1.2: Energy partition in kinetic, potential and surface-pressure terms at (a) initial and (b) final time, with initial-time total energy normalized as one.

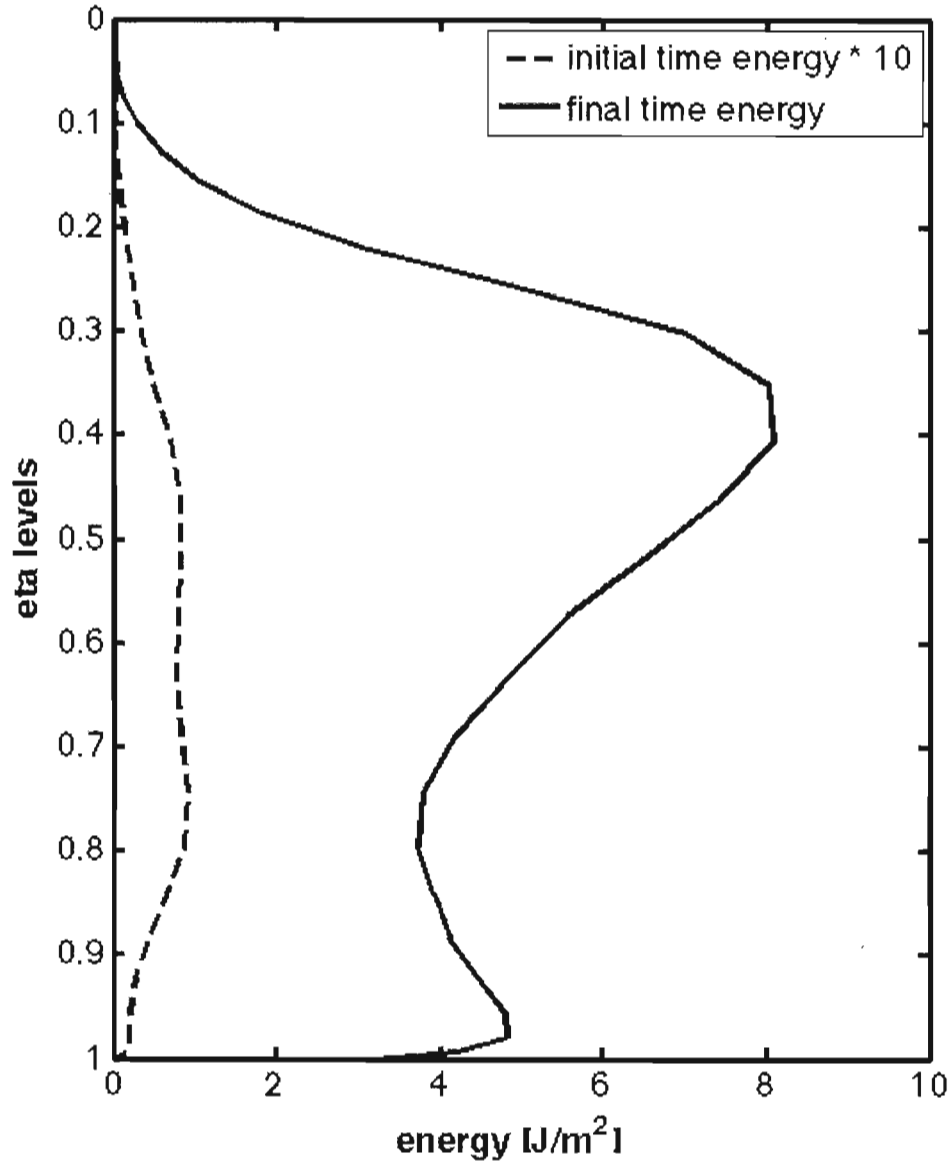
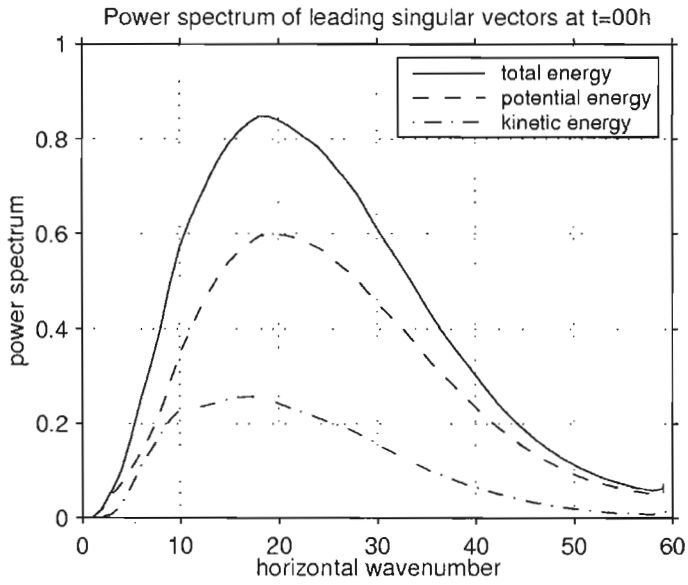


Figure 1.3: Vertical distribution of the total energy averaged among the first ten SVs. For clarity, the initial-time total energy has been multiplied by 10.

a)



b)

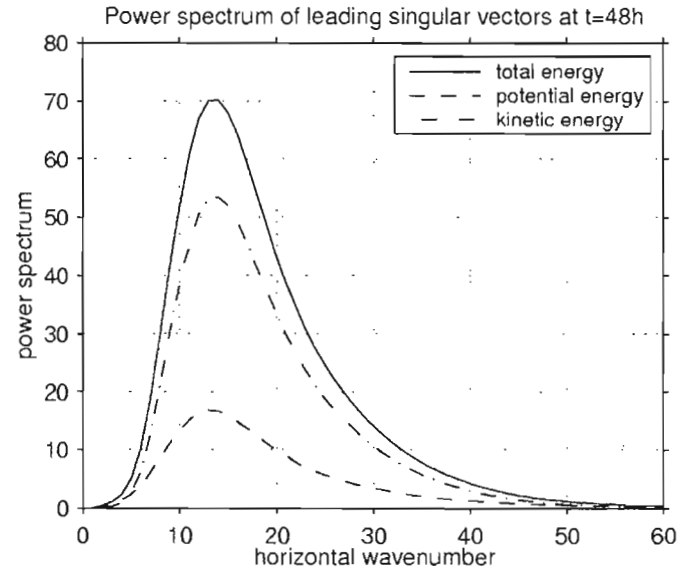


Figure 1.4: The potential, kinetic and total energy spectra at (a) initial and (b) final times, averaged among the first 45 SVs. [Source: Zadra et al. (2004); © Crown copyright, 2004.]

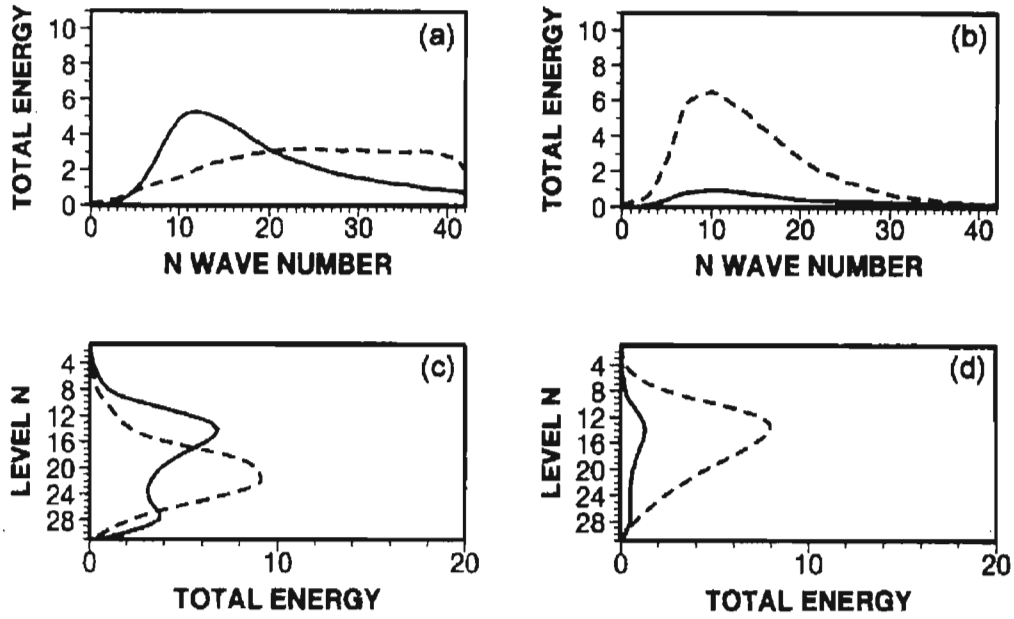


Figure 1.5: (a and b) Spectrum and (c and d) vertical distribution of the total energy for the (a and c) total-energy SVs and (b and d) Hessian SVs. Values at initial (final) time are given by dashed (solid) lines. The total energy at initial time has been multiplied by 100. [Source: Barkmeijer et al. (1999); © Crown copyright, 1999.]

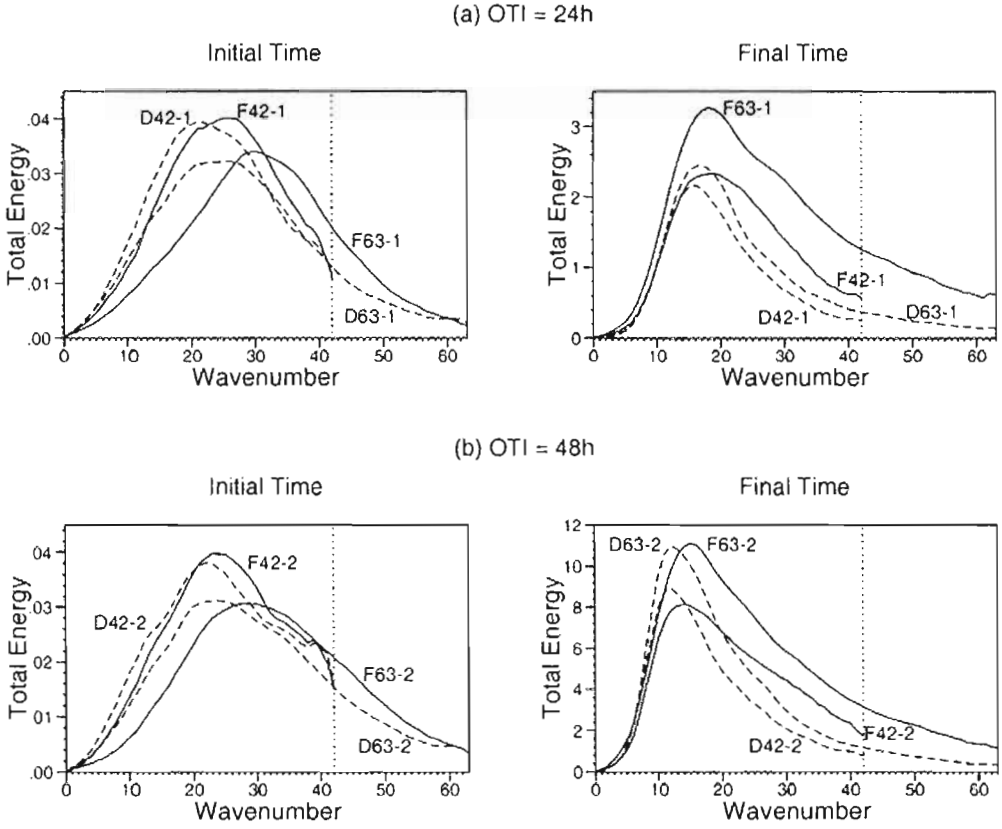


Figure 1.6: The total energy spectrum averaged among the first 10 SVs for experiments with dry physics (dashed lines) and full physics (solid lines) at initial time (left) and final time (right) at OTI of (a) 24 h and (b) 48 h. [Source: Coutinho et al. (2004); © 2004 American Meteorological Society.]

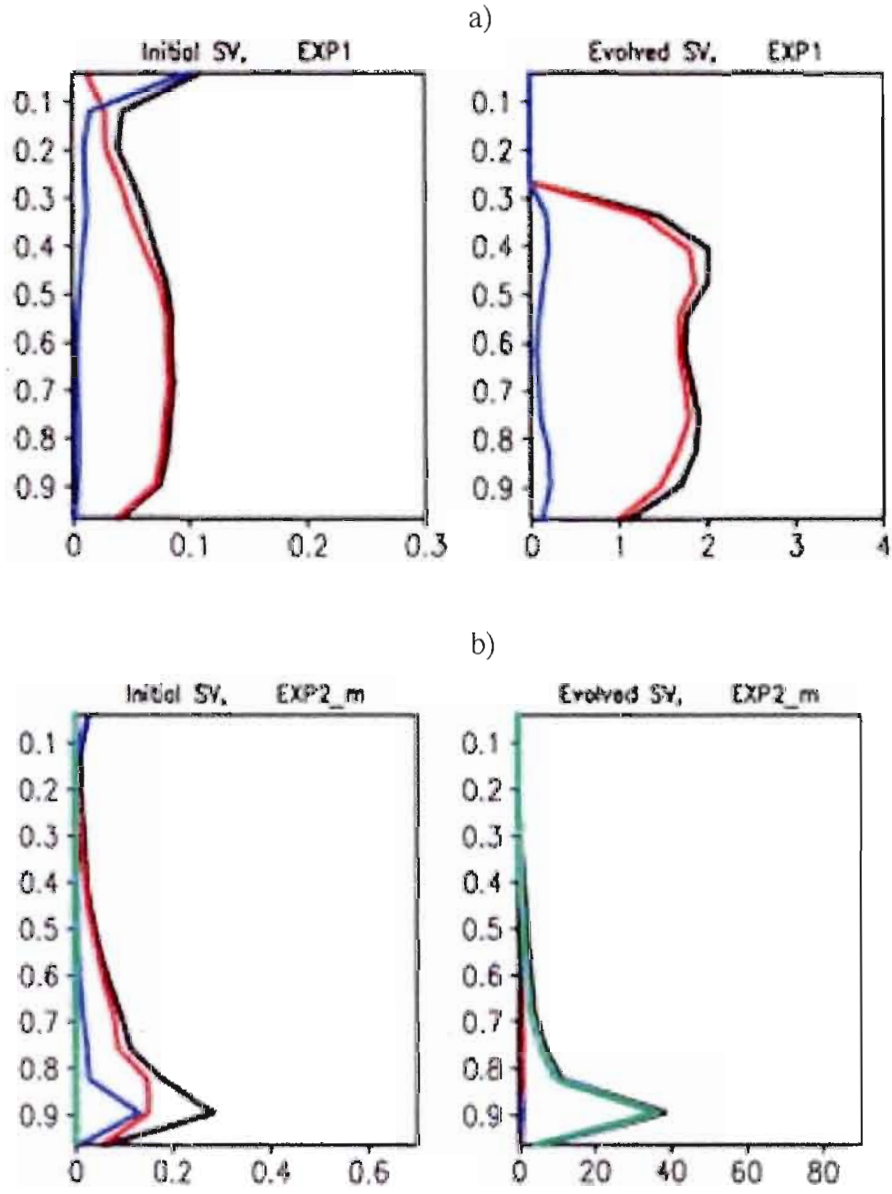


Figure 1.7: Initial and evolved leading SV energy vertical distribution (J/kg) for (a) dry TLM with dry total energy norm and (b) moist TLM with moist total energy norm with full weighting (ε=1). The total energy is represented in black, kinetic energy in red, potential energy in blue and the moist energy in cyan. [Source: Kim and Jung (2009a); © 2009 American Meteorological Society.]

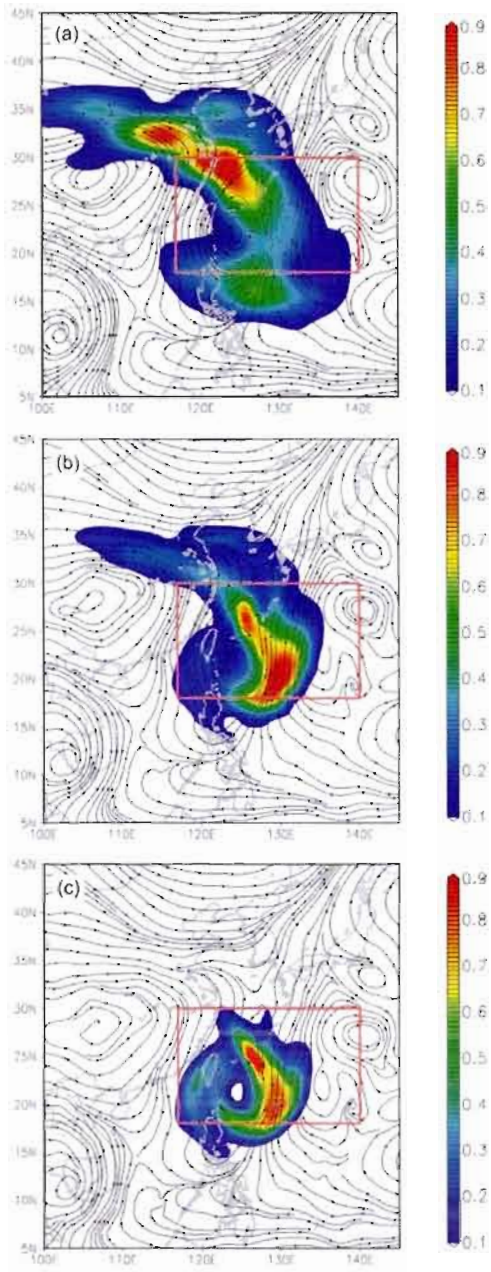


Figure 1.8: Vertically integrated total energy for the initial-time leading SV superimposed on the analyzed streamlines at 500 hPa. The SVs are computed with JMA TL/AD model at a resolution of (a) TL63L60, (b) TL95L60 and (c) TL159L60. The red rectangle represents the target area. [Source: Kim and Jung (2009a); ©2010, the Meteorological Society of Japan.]

CHAPITRE II

LA DÉCOMPOSITION EN VECTEURS SINGULIERS DE LA VARIABILITÉ INTERNE DU MODÈLE RÉGIONAL CANADIEN DU CLIMAT.

Ce chapitre, rédigé en anglais, est présenté sous la forme d'un article qui a été soumis pour publication dans la revue *Climate Dynamics*.

L'article analyse la VI dans des simulations du MRCC effectuées sur un domaine centré sur le continent nord-américain en hiver. L'objectif principal est de déterminer dans quelle mesure la croissance de la VI est expliquée par les perturbations avec le plus fort taux de croissance représentées par les VS. L'analyse porte sur un épisode marqué par une forte augmentation de la VI et une série de dix VS qui sont calculés à partir de l'état de référence du MRCC en utilisant la norme de l'énergie totale localisée sur le domaine d'intégration du MRCC. La norme de l'énergie totale est souvent utilisée pour calculer les VS dans les régions extratropicales parce qu'elle permet de capter la dynamique des instabilités baroclines et barotropes qui se développent dans ces régions. Tel que précisé dans l'Introduction, l'objectif général de la thèse est de vérifier si la croissance épisodique de la VI dans les simulations du MRCC en hiver est due à de telles instabilités. Étant donné que l'analyse porte sur une période hivernale, nous avons décidé d'utiliser un modèle linéaire tangent sec et un intervalle d'optimisation de 36 heures.

L'article offre une analyse comparative entre les caractéristiques de la VI et celles des VS. Il présente aussi le pourcentage dans lequel la VI du MRCC est expliquée par le premier VS ainsi que par la série des dix VS.

SINGULAR VECTOR DECOMPOSITION OF THE INTERNAL VARIABILITY OF THE CANADIAN REGIONAL CLIMATE MODEL

by

Emilia Paula Diaconescu [*,1,3], René Laprise [1,3] and Ayrton Zadra [1,2,3]

[1] *Canadian Network for Regional Climate Modelling and Diagnostics, University of Quebec at Montreal, Montreal, Canada*

[2] *Meteorological Research Division, Environment Canada, Montreal, Canada*

[3] *Centre ESCER (Étude et Simulation du Climat à l'Échelle Régionale)*

Accepted subject to minor modifications for publication in

Climate Dynamics

August 2011

[*] Corresponding author: Emilia Paula Diaconescu
Department of Earth and Atmospheric Sciences
Université du Québec à Montréal (UQAM)
P.O. Box 8888, Succ. Centre-ville
Montréal, Québec
CANADA H3C 3P8

Abstract

Previous studies have shown that Regional Climate Models (RCM) internal variability (IV) fluctuates in time depending on synoptic events. This study focuses on the physical understanding of episodes with rapid growth of IV. An ensemble of 21 simulations, differing only in their initial conditions, was run over North America using version 5 of the Canadian RCM (CRCM). The IV is quantified in terms of energy of CRCM perturbations with respect to a reference simulation. The working hypothesis is that IV is arising through rapidly growing perturbations developed in dynamically unstable regions. If indeed IV is triggered by the growth of unstable perturbations, a large proportion of the CRCM perturbations must project onto the most unstable singular vectors (SVs). A set of ten SVs was computed to identify the orthogonal set of perturbations that provide the maximum growth with respect to the dry total-energy norm during the course of the CRCM ensemble of simulations. CRCM perturbations were then projected onto the subspace of SVs. The analysis of one episode of rapid growth of IV is presented in detail. It is shown that a large part of the IV growth is explained by initially small-amplitude unstable perturbations represented by the ten leading SVs, the SV subspace accounting for over 70% of the CRCM IV growth in 36 hours. The projection on the leading SV at final time is greater than the projection on the remaining SVs and there is a high similarity between the CRCM perturbations and the leading SV after 24 to 36 hours tangent-linear model integration. The vertical structure of perturbations revealed that the baroclinic conversion is the dominant process in IV growth for this particular episode.

Keywords: Internal variability, regional climate model, singular vectors, baroclinic instability.

2.1 Introduction

It is common knowledge that the climate system is a complex dynamical system with natural variability occurring on a wide range of temporal and spatial scales. The complex interactions between and within the different components of the climate system result in atmospheric disturbances with several temporal and spatial scales. While low-frequency fluctuations are induced in the atmosphere by similar fluctuations in the sea surface temperature, soil moisture or interactions with solar radiation, the high-frequency weather fluctuations are associated with nonlinear interactions within the atmosphere. Nonlinearity is a key ingredient of chaotic systems. An important characteristic of chaotic systems is their sensitivity to small differences in initial conditions. This feature has dramatic consequences on the system's predictability, as it will be further explained.

In order to determine the future state of the atmosphere, we use atmospheric climate models, which are mathematical and numerical simplified versions of the atmosphere that either operate on the entire globe (Atmospheric General Circulation Model – AGCM) or on limited regions of the world (Regional Climate Model – RCM). There are two inevitable problems with climate models. First, they are only approximations of the real climate system and therefore are never perfect. Second, these models require initial conditions (ICs) in order to start the simulations. The variables used to simulate the principal atmospheric processes are the temperature, wind, pressure and humidity, and it is impossible to know their exact values at each point in the atmosphere. Therefore the model's ICs always have errors. These two problems associated with the atmosphere's chaotic nature make it impossible to determine exactly the future state of the system beyond a few days, because initially small errors will grow causing the high-frequency part of the model solution to eventually differ from the real atmospheric variations (Lorenz, 1963, 1965). Hence, running a model several times with only small differences in the ICs (reflecting the degree of uncertainty in analysis which is based on observations) will result in different solutions for the same prediction problem. The predictability of synoptic events is lost after a period of several days. In other words, the memory of ICs is lost: any one member become uncorrelated with another member in the same ensemble.

The dispersion of model's solutions caused by its sensitivity to small differences in ICs is named internal variability (IV). IV can be studied by constructing ensembles of simulations where everything (including boundary forcing) is the same, except ICs.

In numerical weather prediction (NWP), the interest is then to provide as accurate as possible information for the ICs. In order to address this issue, data assimilation techniques are used and large ensembles are specifically designed to sample the subspace of the potential errors in observations. Ensemble prediction systems (EPS) have been implemented operationally in the early 1990s at the National Centers for Environmental Prediction (NCEP; Toth and Kalnay, 1997) and at the European Centre for Medium-Range Weather Forecasts (ECMWF; Molteni *et al.,* 1996), and later in other centres such as at the Meteorological Service of Canada (MSC; Pellerin *et al.,* 2003); see Descamps and Talagrand (2007) for an overview. Nowadays, there are four popular methods used for initialization of EPS: the methods of singular vectors (SV), bred modes, ensemble Kalman filter and ensemble transform Kalman filter (these last two methods being ensemble-based data assimilation approaches).

As stated before, the time evolution of any perturbations, no matter how weak they are, will eventually erase the memory of the initial values. Hence, at climate scales, the model's response to initial atmospheric perturbations is independent of the magnitude as well as the type of the perturbation. This behaviour has been verified and confirmed by Lucas-Picher *et al.* (2008a) who initialized the different members of a RCM ensemble by either varying the initialization time between the different members or by adding random or fixed perturbations in some of the atmospheric fields. For AGCM and RCM ensembles, it does not matter how the atmospheric ICs are perturbed. Usually the ensemble is constructed by offsetting the initialization time of each simulation by 6 or 24 hours. After a spin-up period, the IV in AGCM will reach the level of natural climate variability. The behaviour of nested RCM contrasts with that of global models, because the model lateral boundaries (MLB) exert some control on the perturbations' growth and will thus limit the IV. Vukicevic and Paegle (1989) showed that the one-way interacting boundary conditions enhance the predictability of the flow in a local region compared to the case without lateral boundary contraint. They pointed out that the degree of boundary constraint is a function of the size of the domain and

the nature of the flow in the domain.

The issue of the IV in RCM has been investigated previously in several studies (e.g., Giorgi and Bi, 2000; Rinke and Dethloff, 2000; Christensen *et al.*, 2001; Caya and Biner, 2004; Rinke *et al.*, 2004; Vanvyve *et al.*, 2007; Alexandru *et al.*, 2007; Lucas-Picher *et al.*, 2008a; Šeparović *et al.*, 2008; Rapaić *et al.*, 2010). These studies point to the fact that the RCMs' IV depends on the domain's size and location, the atmospheric variable under study, and on the season, being conditioned by the synoptic situation. Lucas-Picher et al. (2008a) have studied the IV for a large domain covering the North American continent. They found that the IV of mean-sea-level pressure and screen temperature exhibits a weak annual cycle with a maximum occurring during spring, while the precipitation IV shows a stronger annual cycle peaking during the summer. In their study over a circumpolar Arctic domain, Rinke *et al.* (2004) found a different annual cycle for temperature with maxima in autumn and winter.

Another important feature of RCM IV is its dependence on the domain size. This characteristic was clearly documented in the study of Alexandru *et al.* (2007), who have compared the IV of an RCM for five different domain sizes on the East Coast of North America during summer of 1993. They found that the IV is weaker for a small domain than for a larger domain, the RCM solution being more strongly constrained by the MLB. Rinke and Dethloff (2000) and Rapaić *et al.* (2010) arrived at the same conclusion when studying the IV for different size domains over the Arctic and North Atlantic regions, respectively.

Rapaić *et al.* (2010) computed the IV separately for the large- and small-scale components of the spectrum, using a length scale of 1400 km for separation. They showed that, for the temperature field, the IV is weaker for the large scales than for the small scales, and noted that the IV for precipitation, a variable that is not strongly constrained by the MLB, is much stronger than for temperature.

An additional factor that can influence the IV is the time an air parcel spends inside the RCM domain, with longer residence times favouring perturbation growth. In their study, Lucas-Picher *et al.* (2008b) showed that the IV for fields that are driven at the MLB (such as mean-sea-level pressure and screen temperature fields) was well correlated with the residence time. On the other hand, the correlation was very weak for precipitation, which is a more local process and is not driven at the MLB. Nevertheless, the strength of the mean flow

influences the divergence of the solutions, through the time required for the perturbations to approach the MLB and exit the domain. As a consequence, the geographical location of the RCM domain also influences the IV. A configuration with a strong through flow will quickly drive the perturbations out of the domain and therefore the growth will be limited; on the other hand, episodes with recirculation flow will lead to high IV (e.g., Laprise *et al.* 2008). Rinke *et al.* (2004) found that the IV is substantially larger for a circumpolar Arctic domain than for middle-latitude domains of comparable sizes; they hypothesised that this is because the Arctic domain is characterized, especially in winter and autumn, by a weaker circulation through the MLB and by an axisymmetric vortex that impedes the perturbation from exiting the domain.

Nikiema and Laprise (2010) have established prognostic budget equations of the IV for the potential temperature and the relative vorticity fields in simulations with the Canadian RCM over a domain centred on the East Coast of North America, for summer 1993. The two equations present similar terms, notably terms relating to the transport of IV by ensemble-mean flow and to the covariance of fluctuations acting on the gradient of the ensemble-mean state. It was noted that episodes of large IV growth usually occur where and when the ensemble-mean state is cyclonic. The analysis of the different components of the IV tendency for an episode with important IV in July 1993 revealed that, for that particular summer episode, the most important contribution to the potential temperature IV growth came from covariance of potential temperature and diabatic heating from convection and condensation. The fluctuation available potential energy, generated by condensation and convection processes, was next converted into fluctuation kinetic energy. For the relative vorticity IV tendency, the most important terms were associated with horizontal components, emphasizing the importance of the horizontal flow in the vorticity IV tendency of this particular episode of high IV. Their study was focused on a summer case of high IV associated with a high synoptic activity. We hypothesize that the results may be different if the analysis is focused on a winter case of IV growth.

All these studies point to the fact that the IV does not have a long-term trend, but is rather characterized by episodes of rapid growth and decay depending on the synoptic situation. The divergence of the RCM solutions appears to be the result of the competition

between two processes. On one side there are perturbations that are excited by the instabilities of the climate system. On the other side there is the forcing exerted by the lateral boundaries of the nested model that encourages the simulations of the RCM toward the same solution; the force of this forcing is given by the strength of the mean flow through the domain, which is strongly dependent on location, season, and domain size. Finally, everything is modulated by the synoptic conditions: periods with important synoptic activity will favour divergence growth, while calm synoptic periods will be characterized by a convergence of the solutions.

In spite of the many studies dedicated to IV, a very small number of them are focused on the physical reasons of its development. It is of great interest to elucidate the dynamical processes responsible for the episodic growth and decay of IV, especially in wintertime when the MLB exercise an important forcing. In this paper we focus on the physical understanding of episodes of rapid growth of IV in simulations of the Canadian Regional Climate Model (CRCM) for December 1992. Our working hypothesis is that IV growth results from rapidly growing perturbations through hydrodynamical instabilities within the regional domain. As a consequence, in periods with large instability the atmosphere should behave more chaotically and small differences would amplify rapidly, while in calm synoptic periods the ensemble members should stay within a narrow range and the IV would have smaller values.

A possible way to approach the issue of hydrodynamical instability is through the singular vector (SV) technique. The SVs from a linearized model provide the most rapidly growing perturbations over a chosen time interval with respect to a given norm. SVs have been used in various applications such as observation targeting (e.g., Buizza and Montani, 1999; Barkmeijer *et al.,* 2001; Wu *et al.,* 2009) and especially in ensemble prediction for generating ICs for NWP EPS. By sampling the phase space in the most unstable directions, the ensemble is supposed to account for a maximum fraction of pdf variance and give to sufficient spread for the ensemble. (e.g., Molteni *et al.,* 1996; Buizza *et al.*, 2008). Further background on SVs is given in Section 2.2.

Numerous studies have emphasized the capability of SVs to capture atmospheric regions and periods characterized by large hydrodynamical instability (e.g., Borges and Hartmann, 1992; Palmer *et al.,* 1994). Our study aims at evaluating whether episodes of

important IV growth can be characterised in terms of SVs. In order to address this issue, a CRCM ensemble of simulations is considered, made of a reference simulation and a set of perturbed member simulations from the reference. We aim at providing a systematic comparison between the CRCM perturbations and the SVs computed using the atmospheric conditions taken from the reference simulation during an episode of important IV growth. Given the linear nature of SVs, this comparison will also provide information concerning the relative roles played by linear and nonlinear processes in the IV growth.

In the following we will first present some general aspects regarding the SV technique (Section 2.2). We will then present the experiment setup with short descriptions of the models in Section 2.3. The main results concerning the CRCM IV, the computation of CRCM perturbations and their comparison with one set of SVs will be presented in Section 2.4. Finally, Section 2.5 will summarize the main findings.

2.2 Singular Vector technique

The SV theory rests on the assumption that the evolution of small perturbations over a short time interval can be described to a good degree of approximation by a linearized system, named the tangent-linear model (TLM).

Consider **X**(t) the state vector, which includes fields such as temperature, wind and surface pressure, and assume that its evolution in time can be described by the equation

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X}). \tag{2.1}$$

where **F**(**X**) represents the nonlinear model tendency. According to the linear theory, the evolution of a small perturbation *x(t)* from the state vector may be approximately described by

$$\frac{dx}{dt} = \mathbf{A}_F x, \tag{2.2}$$

where $\mathbf{A}_F \equiv \left.\frac{d\mathbf{F}}{d\mathbf{X}}\right|_{\mathbf{X}(t)}$ represents the TLM relative to the trajectory **X**(t). Hence, we can use the TLM to describe the time evolution of infinitesimal-amplitude perturbations

superimposed on a fully developed, nonlinear state of a system. The integration of TLM, named the propagator of the TLM, $\mathbf{L}(t,t_0)$, maps initially small perturbations $x(t_0)$ from time t_0 (henceforth referred to as the initial time) to a perturbation at later time t (referred to as the final time) ($x(t)$):

$$x(t) = \mathbf{L}(t,t_0)x(t_0). \tag{2.3}$$

Because it is linearized around the detailed flow from t_0 to t, the TLM depends on the solution of the nonlinear model (the basic state), but it does not depend on the perturbation (Kalnay, 2002). Perturbation growth can be quantified by calculating its amplitude at initial and final times measured by a chosen norm:

$$\begin{aligned} \|x(t_0)\|_{E_0}^2 &= < x(t_0); E_0 x(t_0) > \\ \|x(t)\|_{E_t}^2 &= < x(t); E_t x(t) > \end{aligned} \tag{2.4}$$

Here, < , > denotes the Euclidean inner product, E_0 represents the initial-time norm and E_t the final-time norm.

SVs are defined as the perturbations that, under linearized dynamics about a time-evolving basic flow, grow most rapidly over a given time interval known as "optimization time interval", according to a given norm. SVs can be obtained by solving the generalized eigenvalue problem

$$\mathbf{L}^* E_t \mathbf{L} y_i(t_0) = \lambda_i^{\,2} E_0 y_i(t_0), \tag{2.5}$$

where $\mathbf{L}^*$ is the adjoint of the operator $\mathbf{L}$, λ is the singular value, and $y_i(t_0)$ is the initial-time SV. Hence, the problem of finding the perturbations with maximum growth in terms of a norm E, that is the singular vectors, can be reduced to the search of the eigenvectors of the matrix $\mathbf{E}_0^{-1/2}\mathbf{L}^* E_t \mathbf{L}\mathbf{E}_0^{-1/2}$ with the largest eigenvalues λ^2. Note that the left-hand side of the equation involves one integration with the TLM, followed by one integration with the adjoint model. The equation can be solved efficiently using software packages such as the Arnoldi Package (ARPACK).

The resulting SVs form an orthogonal set and can be ordered according to growth rate, with the fastest growing structure being the first singular vector. The first SV, also referred to

as the leading SV and noted *SV1*, is the SV that maximizes the amplification factor defined as the ratio between the final and the initial norms. This ratio is equal to the square of the singular value:

$$\frac{\|x(t)\|^2}{\|x(t_0)\|^2} = \frac{< x(t), E_t x(t) >}{< x(t_0), E_0 x(t_0) >} = \frac{x^T(t_0)\mathbf{L}^* E_t \mathbf{L} x(t_0)}{x^T(t_0) E_0 x(t_0)} = \lambda^2. \tag{2.6}$$

Once the initial-time SVs are obtained, the corresponding final-time SVs can be derived by integrating the TLM.

The SVs associated with the largest λ are also called leading SVs. Given an orthogonal and complete set of SVs ordered according to amplification factor, any initial perturbation can be decomposed in terms of the complete set of SVs comprising growing, neutral and decaying solutions:

$$x(t_0) = \sum_{j=1}^{N} \alpha_j \hat{y}_j(t_0), \tag{2.7}$$

where $\hat{y}_j(t_0)$ are the initial-time SVs normalized to have the unity norm, and $\alpha_j = < x(t_0), E_0 \hat{y}_j(t_0) >$ are the respective projection coefficients. Because the singular vectors are orthonormal, the amplitude of the initial-time perturbation is

$$|x(t_0)| = \sqrt{\|x(t_0)\|^2} = \sqrt{\sum_{j=1}^{N} \alpha_j^2}, \tag{2.8}$$

and its amplitude at time t is

$$|x(t)| = \sqrt{\|x(t)\|^2} = \sqrt{\sum_{j=1}^{N} \alpha_j^2 \lambda_j^2}. \tag{2.9}$$

Therefore the final-time magnitude of the perturbation depends on the projection coefficients α_j^2 weighted by the individual amplification factors λ_j^2. Hence, growth is expected only if the magnitudes of the projections on the growing SVs ($\lambda_j^2 > 1$) are

sufficiently large with respect to those for decaying SVs (λ_j^2<1) (Ehrendorfer and Errico, 1995; Errico *et al.,* 2001).

The dimension of the model gives the number of SVs. However, in the most part of previous studies, truncated bases of SVs are used instead of the complete base. In the particular case of a growing perturbation, and for a truncated base with a small number of SVs, if only a few singular values are large, it is expected that the projection onto the leading SVs will dominate (Hartman *et al.,* 1995). On the other hand, if the number of SVs is very large, the overall projection onto the slow growing and decaying SVs could overwhelm the projection onto the leading SVs (Errico *et al.,* 2001). For example, Errico *et al.* (2001) have used 13 sets with 100 SVs. They found that only small fractions (approximately 1%) of the possible number of SVs are fast growing structures; the others are very slowly growing structures. In our analysis, we make the hypothesis that the CRCM ensemble perturbations, in periods of growing IV, are fast-growing perturbations that may be efficiently described by a subset of the 10 leading SVs, the remaining slowly-growing SVs being of little interest in our case.

2.3 Experimental set-up

The IV of an RCM is studied through an ensemble approach. An ensemble of simulations performed with a single version of a regional model with different initialization times and driven by the same set of lateral boundary conditions can be considered as a sample of the RCM solution space (e.g., Christensen *et al.,* 2001).

2.3.1 The Canadian Regional Climate Model ensemble.

In the present work, the IV is explored by constructing an ensemble of 21 integrations with version 5 of the Canadian Regional Climate Model (CRCM_5), developed by the Canadian Regional Climate Modelling and Diagnostics Network in collaboration with the Meteorological Service of Canada (MSC) (Zadra *et al.,* 2008). CRCM_5 is a limited-area version of the Global Environmental Multiscale (GEM) model employed for numerical weather prediction at the MSC (Côté *et al.,* 1998; Yeh *et al.,* 2002). The model solves the

fully elastic nonhydrostatic equations with a hybrid vertical coordinate based on terrain-following normalized hydrostatic pressure (Laprise, 1992), $\eta = (p - p_{top})/(p_s - p_{top})$, where p_{top} and p_s are the pressure at the model top and at the surface, respectively. However, in the configuration we used, the model was run in hydrostatic mode. In our experiments, the model includes 53 irregularly spaced hybrid levels in the vertical between the surface and 10 hPa, and a limited-area, rotated, latitude-longitude mesh with 120 x 120 grid points with spacing of 0.5 degree, or about 55 km. The time step is 30 minutes. A ten grid-point wide Davies-type lateral boundary relaxation zone (Davies, 1976) is applied. No large-scale spectral nudging (Alexandru *et al.,* 2009) has been performed in this study. The integration area, shown in Figure 2.1, is centred on the North American continent.

In our configuration, the model uses the Interactions Soil-Biosphere-Atmosphere (ISBA) as land-surface scheme (Bélair *et al.*, 2003), the Kain-Fritsch scheme (Kain and Fritsch, 1990) for deep convective processes, the Kuo transient scheme for shallow convection (Kuo, 1965; Bélair *et al.,* 2005) and the Sundqvist scheme (Sundqvist *et al.,* 1989) for large-scale condensation. The radiation package for solar and terrestrial radiation is based on the correlated-K approach (Li and Barker, 2005). Subgrid-scale orographic gravity-wave drag is due to McFarlane (1987) and low-level orographic blocking is described in Zadra *et al.* (2003).

Ocean surface conditions are prescribed. Sea surface temperatures (SST) and sea-ice are interpolated from the Atmosphere Model Intercomparison Project (AMIP2; Gleckler, 1996) available on a one-degree latitude-longitude grid for monthly mean values.

Initial and lateral boundary conditions have been taken from the ECMWF (European Centre for Mèdium-range Weather Forecasts) reanalysis project ERA40 (Uppala *et al.,* 2005) available every 6 hours.

Using the ensemble described here, we investigate the IV for December 1992. The 21 simulations have the same set-up except for the ICs that correspond to different starting dates between October 23 to November 12 at 00:00 UTC. The corresponding 21 runs are noted as *NA23*, *NA24*, ... , *NA31*, *NA01*, ... , *NA11*, *NA12*, the number representing the day of the ICs. Model output was archived every six hours for each simulation. Because the simulations use

the same driving fields and the same surface fields in the ICs, which are interactive afterwards, there are no lateral and lower boundary conditions perturbations. By design, the dispersion of model simulations is due only to the differences in the atmospheric initial conditions.

The first episode in the month of December 1992 when the IV begins to grow substantially in our ensemble of simulations occurred on December 4th (Fig. 2.2). Amongst the 21 members, the member named *NA11* was the closest one to the ERA40 driving data; for this reason we chose that member as the CRCM reference simulation for computing the IV.

2.3.2 The SVs set.

The software required for the calculation of the SVs on a limited-area domain was not yet ready at the time of our investigation. Hence the SVs were computed using the tangent-linear (TLM) and adjoint versions of global GEM model, originally developed for 4D variational data assimilation system (Gauthier *et al.*, 2007).

According to Eq. 2.5, several choices must be made when SVs are computed, such as the optimization time interval (OTI) and the initial and final norms. We computed several sets of SVs with different set-ups, but in the present paper we present only one set. Results for the other sets are qualitatively similar. The set presented in this paper has the following characteristics.

Since our analysis is focused on a winter period, we opted for TLM with a simplified physics containing only the vertical diffusion scheme. A detailed description of the available simplified parameterizations and their impact on SVs computed with the total-energy norm can be found in Zadra *et al.* (2004). The calculation of SVs is restricted to perturbations in horizontal wind ($\mathbf{V}=(u,v)$), temperature (T) and surface pressure (p_s) fields.

The evolution trajectory was generated using the nonlinear full-physics global GEM model starting from 4 December 1992 at 12:00 UTC, which represents the moment when the CRCM IV begins to grow. The corresponding IC for the global model were constructed by combining the reference CRCM simulation (*NA11*) over the region encompassed by the CRCM integration, with those of the ERA40 reanalysis elsewhere on the globe.

The nonlinear, tangent-linear and adjoint global models were run with uniform horizontal resolution of 1°, 28 eta levels and a time step of 30 minutes. The OTI is 36 hours, which corresponds to period of initial growth of the CRCM IV. Due to limits of computational cost, the number of SVs was restricted to the leading ten, and output every 6 hours.

As initial and final norms, we opted for the dry total-energy norm restricted to a sub-domain within the CRCM integration area (the region delimited by the red line in Figure 2.1), which permits the selection of SVs situated within the active region of CRCM simulations. The dry total-energy norm is a rather common choice in the computation of singular vectors, and it provides a relative weighting between the mass (T, p_s) and wind (u, v) fields (Buehner and Zadra, 2005). Here, it is defined by

$$\begin{aligned} < x, Ex > = & \iint_A \left[\int (u^2 + v^2) \frac{1}{g} \frac{\partial p}{\partial \eta} d\eta \right] dA \\ & + \iint_A \left[\int \left(\frac{c_p}{T_r} T^2 \right) \frac{1}{g} \frac{\partial p}{\partial \eta} d\eta \right] dA \\ & + \iint_A \left[\frac{R T_r}{p_r g} p_s^2 \right] dA \end{aligned} \qquad (2.10)$$

where A is the horizontal domain, g is the gravity constant (= 9.806 m s^{-2}), R is the gas constant for dry air (= 287.04 J K^{-1} kg^{-1}), c_p is the specific heat for dry air at constant pressure (= 1005.46 J K^{-1} kg^{-1}), T_r is a reference temperature (= 300 K) and p_r is a reference pressure (= 1000 hPa). Here $x = (u, v, T, p_s)$ represents a perturbation of the model state vector comprising horizontal wind, temperature and surface pressure perturbations. The three terms on the right-hand side represent the kinetic, potential and surface-pressure components of the dry total energy, respectively. The vertical integral is restricted to the levels situated between hybrid levels η=1 and η=0.1, the levels above approximately 100 hPa being neglected in order to eliminate potential growing modes, likely numerical in origin, that may be restricted to the top few model levels.

Hence, SVs are available in the form of perturbations in wind, temperature and surface pressure fields every six hours on a grid-spacing of 1° and 28 eta levels. In order to facilitate

the comparison with the CRCM fields, the SVs are interpolated on the 0.5° CRCM grid and all the results are calculated with respect to the dry total-energy norm defined over the CRCM domain. The comparison between CRCM perturbations and SVs is based on their spatial structure, as well as their temporal evolution, partition and distribution of energy. We also compute the projection of CRCM perturbations on the set of SVs, which provides a quantitative measure of their similarity.

2.4 Results

Results are presented in three subsections. The first subsection describes the CRCM IV, the second is focused on the set of SVs and a comparison with the CRCM perturbations, and the last is dedicated to the projection of the CRCM perturbations on the set of SVs.

2.4.1 CRCM internal variability

The IV is defined as the spread between the members in an ensemble with respect to a reference state:

$$\sigma_{IV}^2(x,y,z,t) = \frac{1}{M-1}\sum_{m=1}^{M}\left[x'_m(x,y,z,t)\right]^2 \tag{2.11}$$

Here, M corresponds to the total number of members in the ensemble of simulations, $x'_m(x,y,z,t)$ represents the perturbation of member m (X_m) for fields *(u, v, T, p_s)* function of time and location on the three-dimensional model grid,

$$x'_m = X_m - X_{ref}, \tag{2.12}$$

and X_{ref} represents the reference state. In several other studies, IV has been defined using the ensemble mean (EM) as reference

$$X_{ref} = \overline{X}^M = \frac{1}{M}\sum_{m=1}^{M} X_m \tag{2.13}$$

(e.g., Alexandru *et al.*, 2007; Šeparović *et al.*, 2008). It is important however to realise that the *EM* is in general not a solution of the atmospheric equations of motion. Also, the *EM* lacks the fine scales that tend to characterise the irreproducible components, in the nomenclature of Šeparović *et al.* (2008). Because we will also want to use X_{ref} as the trajectory in the calculation of the SVs, it must correspond to a solution of the equations of motion. In the following we have opted to use one member in the ensemble as reference; the member was chosen as the one closest to the driving conditions during the episode under study. For our case this corresponded to *NA11*, as mentioned earlier.

We mention that, in the case of an AGCM, the spread of an ensemble is very large approaching the value of natural transient variability. As a consequence, the ensemble members will be very different and can have different stability characteristics. However, in the case of a RCM, the spread of an ensemble is limited by the constraint exerted by the model lateral boundaries, which force all members towards the same solution as the driving field in the region of lateral boundaries. Therefore, the RCM members do not have time to diverge very much and will present similar features. The differences between the members of the RCM ensemble can hence be regarded in this case as small perturbations from a time-evolving reference state.

In this study we have chosen to express the IV in terms of energy. For each perturbation $x'_m(u_m, v_m, T_m, p_{sm})$ we have computed the perturbation total energy (E_m) equivalent to the SV dry total-energy norm (Eq. 2.10). The ensemble-average energy of these perturbations provides an equivalent measure for the IV:

$$E_m = \iiint_V \rho\,(u_m{}^2 + v_m{}^2)\,dVol + \iiint_V \rho\,(\frac{c_p}{T_r} T_m{}^2)\,dVol + \iint_A (\frac{RT_r}{p_r g} p_{sm}{}^2)\,dA$$
$$E_n = \frac{1}{M-1} \sum_{\substack{m=1 \\ m \neq n}}^{M} E_m \qquad (2.14)$$

where n indicates the reference member.

Figure 2.2 shows the time evolution of the CRCM IV during the month of December 1992, expressed in terms of the total energy of perturbations. Note the episodic character of the IV, with various episodes of growth resulting in large IV. In this study we focused on the

first maximum reached by 6 December at 12:00 UTC. Additional information is provided by the energy partition into kinetic, potential and surface-pressure terms, their time evolution is also shown in Figure 2.2. During the entire period, most of the energy is in the kinetic form, with the surface-pressure term being relatively small. At about 12:00 UTC on 6 December, both the kinetic and potential energy components exhibit a maximum.

Figure 2.3 displays the horizontal distribution of the vertically integrated energy of the CRCM-perturbations, together with the 500-hPa geopotential field from the reference simulation. On 4 December at 12:00 UTC (Figure 2.3a), some perturbations were present in two distinct regions of the domain: one over central Canada and another one over Québec-Labrador, close to the eastern boundary of the domain. As time progressed, the latter perturbation diminished and eventually disappeared, while the central-domain perturbation kept growing and moved toward the East Coast of the USA (Figure 2.3f). By 00:00 UTC on 6 December (Figure 2.3e), the total energy distribution exhibits a large elongated maximum extending from south of the Great Lakes to the Atlantic Coast, close to the model's eastern lateral boundary. The presence of the boundary inhibits the further development eastward because the driving boundary fields in the atmospheric variables (u, v, T) are the same for all simulations. As a consequence, a gradual decrease in the IV is noted close to the MLB, and the IV is actually zero in the boundaries relaxation zone. The peak in IV intensity is reached by 12:00 UTC 6 December and comes essentially from the maximum located south of the Great Lakes.

2.4.2 CRCM perturbations versus SVs

This section compares the CRCM IV perturbations with a set of ten leading SVs, computed as described in Section 2.2. The initial time for SVs corresponds to 4 December at 12:00 UTC when the CRCM IV begins to grow. In order to avoid the influence of the MLB on CRCM IV, the final time for the computation of the SVs was chosen to be 6 December at 00:00 UTC. The extent of final- and initial-time norms is within the CRCM domain, as shown in Figure 2.1.

a) *Time evolution of energy*

The total-energy evolution of the ten leading SVs during the 36-hour period, 12:00 UTC 4 December – 00:00 UTC 6 December is shown in Figure 2.4. For each SV, the energy was normalized by its initial-time value, i.e. it is set to 1 Jm^{-2} at the initial time. Note that all ten leading SVs are growing perturbations, with the first leading SV having an energy amplification factor greater than 110.

Some details on the perturbation growth can be found in the energy partition during the growth period; this is shown in Figures 2.5 and 2.6. At the initial time, most of the SV total energy is in the form of potential energy while later, kinetic energy dominates; this is a typical feature of baroclinic disturbances that are usually captured by extratropical SVs. Figures 2.6a and 2.6b display the energy partition at initial and final time and show all ten SVs exhibit similar properties. The initial-time energy partition of SVs is different from that of CRCM perturbations. This suggests that the SVs are outside of the attractor at initial time. (Kalnay, 2000). However, in only one step, the dominance of energy components is renversed. Hence, with the exception of the initial time, the other steps are characterised by the dominance of kinetic energy – as in the case of CRCM perturbations (Figure 2.2). This rapid change in the SV energy partition toward the CRCM-perturbation partition can be interpreted as a rapid (one time only) rotation of the initial-time SVs toward the system attractor (e.g., Szunyogh *et al.*, 1997; Kalnay, 2000). By December 6^{th} 00:00 UTC, the ratio of kinetic energy to potential energy for the CRCM perturbations varies between 3.58 (for *NA28*) and 4.14 (for *NA08* and *NA10*). For the set of SVs at final time, the ratio has values varying from 3.3 to 4.7. Also, the time evolution during the growing period shows that the CRCM-perturbation kinetic energy increases by a larger rate than the potential energy, as is the case for SVs.

b) Structure of the leading SV

The horizontal and vertical structures of the leading SV, normalized to have unit energy at initial time as explained earlier, is presented in terms of temperature perturbations in Figure 2.7a-d, meridional wind in Figure 2.8a-d, and zonal wind in Figure 2.9a-d. The structure at initial time is presented in the panels *a)* and *b)*, while the panels *c)* and *d)* show

the evolved structures according to the TLM after 30 hours. At initial time, the leading SV has the structure of a wavepacket located in the centre of the domain, with perturbation maxima of T=0.07 °C, v=0.07 m/s and u=0.04 m/s. After 30 hours, the perturbation has moved southeast and reached the USA East Coast and Atlantic Ocean. The horizontal scale of the wavepacket has since grown considerably, and the amplitude maxima have grown to reach values of T=0.2 °C, v=0.48 m/s and u=0.35 m/s. Note once more that the amplification of the wind field is larger than that of the temperature field.

Regarding the vertical structure, Figures 2.7b, 2.8b and 2.9b reveal at initial time a strong westward (upstream) tilt with height, especially in the mid-troposphere, the structure at the upper levels being situated close to the western limit of the initial-norm domain. As can be seen in Figures 2.7d, 2.8d and 2.9d, the westward tilt reduces as the perturbation grows and only a small tilt remains in the main pattern of the temperature and zonal-wind perturbation, which suggests that the perturbation will continue to grow but at a slower rate. As explained by Hoskins, Buizza and Badger (2000) and Coutinho, Hoskins and Buizza (2004), this westward tilt is a major characteristic of mid-latitude baroclinically growing non-normal modes. This configuration allows the perturbation to grow by a conversion of the basic flow available potential energy into perturbation kinetic energy. The other 10 leading SVs have similar structures and evolutions (not shown).

For comparison, the structure of the perturbations in one of the 20 CRCM members after 30 hours is presented in Figures 2.7e-f for the temperature field, in Figures 2.8e-f for meridional wind and in Figures 2.9e-f for zonal wind. Note the similarity to the SV spatial patterns for all three fields, especially for the maxima situated on the continent. The temperature pattern located over the Atlantic Ocean is very weak in the CRCM case, perhaps due to the vicinity of the eastern MLB of the regional model.

We must mention that the structure of the initial SV is very different from that of CRCM perturbations (not shown) because the initial SVs are outside the attractor, pointing to areas in the system phase space where solutions do not naturally occur. However, the SVs rapidly rotate towards the attractor and after 30 hours the leading SV shape is almost identical to the CRCM evolved perturbation.

c) *Horizontal and vertical distribution of energy*

Figure 2.10 presents the vertically integrated total energy of the CRCM IV calculated with the 20 members on December 5 at 18:00 UTC, and for the leading SV after 30 hours. Note the colocation of the maxima and the similarity of the overall pattern. Nevertheless, the SV is slightly displaced eastward of the CRCM perturbations, while the CRCM field is more extended over the American continent.

The rapid rotation of SVs toward the CRCM solutions is also reveal by Figure 2.11, which displays the vertical distributions of the horizontally integrated kinetic and potential energy of the leading SV and of the CRCM's IV, on December 4th 12:00 UTC (initial time; top panels), 24 hours later (middle panels), and on December 6th 00:00 UTC (low panels). Once again, the SV's amplitude has been normalised to correspond to a unit total energy (1 Jm^{-2}). At initial time, the potential energy of the leading SV is larger than the kinetic component and peaks around the 700-hPa level (right panel, Figure 2.11a). This is quite different from the CRCM, where the kinetic energy dominates, with a maximum near 400 hPa. 24 hours later, the SV kinetic energy is much greater than the potential energy, similar to CRCM perturbations, with maximum kinetic energy near 400 hPa. The similarity persists up to 36 hours. Nevertheless, some differences are noted at lower levels, close to the surface, especially in the potential energy, which has a near-surface maximum in the case of SV (this was already noticeable in Figures 2.7d and 2.7f).

Correlation coefficients between the vertical distribution of CRCM IV energy and of the leading SV as a function of time are shown in Figure 2.12a. For the total energy, the correlation grows from a minimum value of 0.7 at the initial time (December 4th 12:00 UTC) to 0.96 at 24 hours. At this time, all CRCM members perturbations have a similar high correlation coefficient with the leading SV, as can be seen in Figure 2.12b. The correlations are higher when only kinetic energy is considered and much smaller for the potential energy. These small correlations in potential energy are due primarily to the surface maximum noted in the SV case. Several factors can explain the differences at lower levels. First, interpolation/resolution issues: SVs are computed on a global uniform grid with 1° horizontal

resolution and next interpolated on the CRCM grid, which has a 0.5° horizontal resolution. Second, the accuracy of the TLM: the evolved SVs are the result of TLM propagation with simplified physics, while CRCM perturbations are the result of non-linear model integration with complete physics, and at lower levels the non-linear processes can be even more important.

It is interesting to note, however, that the overall comparison between the CRCM perturbations and the set of SVs has showed a remarkable similarity between the CRCM perturbations and the most unstable SV after 30 hours of integration, especially in regions with maximum total energy.

2.4.3 Projection on the set of singular vectors

Let $x_k(u_k, v_k, T_k, p_{Sk}, ...)$ represent the CRCM k^{th} member perturbation, and $y_j(u_j, v_j, T_j, p_{Sj}, ...)$ be the j^{th} SV. Given the orthogonality for the total energy, and if the SV are normalized (total energy equal to one), the projection of x_k on a complete base of N SVs at time t is:

$$x_k(t) = \sum_{j=1}^{N} \alpha_{kj} \hat{y}_j(t), \tag{2.15}$$

where $\hat{y}_j(t)$ represents the total-energy normalized SV and $\alpha_{kj} = < x_k(t), E_t \hat{y}_j(t) >$ the projection coefficients. If a truncated (not complete) basis of ten SVs is considered, the CRCM perturbation can be written as:

$$x_k = \sum_{j=1}^{10} \alpha_{jk} \hat{y}_j + \sum_{j=11}^{N} \alpha_{jk} \hat{y}_j = \tilde{x}_k + \Delta x_k, \tag{2.16}$$

with $\tilde{x}_k$ the projected part into the truncated basis, and Δx_k the non-projected part representing slower growing, neutral or decaying vectors.

The total energy of the projected part will be given by the sum of the squares of the projection coefficients,

$$E(\tilde{x}_k) = \left\|\tilde{x}_k(t)\right\|^2 = \sum_{j=1}^{N} \alpha^2{}_{kj}\,, \tag{2.17}$$

and the CRCM perturbation total energy can be written as:

$$E(x_k) = \sum_{j=1}^{10} \alpha_{jk}{}^2 + \sum_{j=11}^{n} \alpha_{jk}{}^2 = E(\tilde{x}_k) + E(\Delta x_k). \tag{2.18}$$

Figure 2.13a displays the average values of the energy corresponding to all the CRCM members (*NA23* to *NA12*): the total energy ($\frac{1}{20}\sum_{k=1}^{20} E(x_k)$) in a solid black line, and the non-projected part of total energy ($\frac{1}{20}\sum_{k=1}^{20} E(\Delta x_k)$) in a dashed black line. The red area indicates the projected part of total energy ($\frac{1}{20}\sum_{k=1}^{20} E(\tilde{x}_k)$), i.e. the part of CRCM perturbations' total energy represented by the ten leading SVs. The particular case of one CRCM member, *NA28*, is displayed in Figure 2.13b. Note that initially only a negligible part of the CRCM perturbations are projected into the ten SVs sustaining the idea that SVs are ouside the system's atractor at initial time. As time passes, the projected part grows rapidly, and after 36 hours it represents a significant fraction of the total CRCM perturbation growth: by 00:00 UTC 6 December, the projected part reaches the value of 1.35 x 10^5 J m^{-2}, which is about 40% of the average total energy of the CRCM perturbations, or 70% of the CRCM IV growth in 36 hours, the other 30% of the CRCM growth being due to the non-projected part growth. Therefore, 70% of the CRCM-perturbations growth is explained by the first ten SVs. This represents an average over the twenty CRCM members. In the case of *NA28* member (Figure 2.13b), the non-projected part diminishes in time and consequently the projected part accounts for the entire growth.

Figure 2.14a and 2.14b show the 6h-mean growth rate of the projected and the non-projected parts, estimated as

$$r_k(t) = \frac{1}{\Delta t}\frac{E_k(t) - E_k(t-\Delta t)}{E_k(t-\Delta t)}, \tag{2.19}$$

with $\Delta t = 6h$, and expressed in units of percent per hour. Figure 2.14a represents the perturbations of all CRCM members (shaded area) and their mean value (the solid line) while Figure 2.14b represents the case of one CRCM member perturbation (*NA28*). During the first hours, the projected part (the part represented by the first ten SVs) is made of small-scale small-amplitude perturbations with large growth rates of 80% to 170% per hour, while the non-projected part is composed of large-scale large-amplitude perturbations with very small growth or even decaying rates as in the case of *NA28* perturbation (Figure 2.14b). As time passes, the projected part grows in spatial scale and gains in amplitude, but its growth rate decreases, approaching the rate of the non-projected part.

The rotation of leading SV toward the CRCM solutions is also evident in Figure 2.15, which shows the meridional wind near 460 hPa of the *NA08* pertubation projected on the leading SV (Figures 2.15a and 2.15b) and of the *NA08* total perturbation (Figures 2.15c and 2.15d) at initial time and after 30 hours. The difference between Figure 2.15a and Figure 2.15c illustrates the relatively large scales aspect of the non-projected part at the initial time and the very small projection of the CRCM perturbation on the initial SV..

As mentioned previously, the first leading SV distinguishes itself from the other SVs by its high amplification rate. An important question that arises from this is whether there is a preference of the projection of CRCM IV on this particular SV. Figure 2.16 shows fraction of CRCM perturbations energy explained by each of the ten SVs at 18:00 UTC on December 5th and confirms the perturbations project significantly on the first leading SV. The first leading SV grows rapidly enough to account in average for over 23% of the CRCM total perturbation energy and for more than 50% of the projected part. Note that a particular member, *NA12*, has a very small projection. This is due to the fact that the NA12 is very close to the reference simulation; as a concequence, the perturbation energy of this particular member is very small, its total energy being about eight times smaller than the average total perturbation energy of other members in the ensemble.

While, Figure 2.16 presents the fraction of each of the CRCM-perturbation energy explained by each of the ten SVs, Figure 2.17 shows the average value (ensemble mean) of the CRCM projected part on each of the ten SVs, from the initial time to the final time. The figure reveals that the perturbation is predominantly projected on the most unstable SV

through most of the growth period. At final time, the projections on each of the other nine SV are small compared to the projection on the leading SV. However, when adding them together, they have an important contribution to the total projected part.

The set of SVs used in this study contains only the first ten SVs. Studies that used more than ten SVs (e.g., Errico *et al.,* 2001; Snyder and Hakim, 2005) have shown that if the number of amplifying SVs is very large and the spectrum of growth rates is flat, then the probability that a random perturbation projects strongly on any single or small set of SVs is correspondingly small. Even though subsequent growth in the leading SVs will be greater than growth in other individual SV, the overall behaviour can be dominated by the much larger set of non-leading SVs.

2.5 Summary and concluding remarks

This paper focused on the physical understanding of specific episodes with rapid growth of IV in Canadian Regional Climate Model (CRCM) simulations. It was hypothesized that periods of important IV growth might arise from local dynamical instabilities. Initially small perturbations in unstable regions begin to develop while they are advected by the mean flow toward the lateral boundaries where they are eventually transported out of the domain.

A possible way to approach the question of hydrodynamical instabilities is through the singular vector (SV) approach. Hence, our focus was on comparing the CRCM perturbations with a set of SVs and ascertaining whether the IV growth may be linked to the most unstable SV.

To test the hypothesis, a 21-member CRCM ensemble of simulations was performed and the analysis was focused on one specific episode of large IV growth occurring between 4 and 6 December 1992. The ensemble was decomposed into a reference simulation and a set of perturbations with respect to that reference. Then, the total energy of each CRCM perturbation was computed, equivalent to the dry total-energy norm used to compute the set of SVs. The set of SVs was restricted to the leading ten, archived at every 6 hours, using an optimization time interval (OTI) of 36 hours, and initial- and final-time norms restricted to a

sub-domain contained within the CRCM integration area. We found that the ten leading SVs were all growing perturbations, the first SV dominating the others.

The comparison was first focused on SV properties such as perturbation structure, temporal evolution of energy, energy partition and spatial distribution. A remarkable structural similarity was found between the CRCM perturbations and the leading SV after 24- to 36-hour tangent-linear model integration. The spatial patterns were roughly the same, presenting a westward (upstream) tilt in the temperature and zonal wind fields. It was also shown that all ten SVs present at final time a similar partition of energy, with dominant kinetic energy component, as for the CRCM perturbations. The overall agreement was illustrated by comparing the horizontal and vertical distribution of the first SV's energy with the horizontal and vertical distribution for the average of the energy of the CRCM perturbations. The comparison revealed similar overall patterns and a colocation of total-energy maxima after 24 to 36 hours.

We have also computed the projection of the CRCM perturbations on the truncated base of ten leading SVs. It was shown that only a very small part of the CRCM perturbations initially projected on the ten SVs. During the next 36 hours, the projected part grew very much, and ended up representing an important part of the total CRCM perturbation growth after 36 hours. Quantitatively, up to 40% of the average CRCM total energy was projected on the ten SVs at final time, the first SV accounting in average for over 23% and the rest being projected into the nine remaining SVs, for which only small projection amplitudes did occur. Other slowly growing or decaying perturbations represented the non-projected part. As a consequence, the overall picture showed that even though subsequent growth in the leading SV was greater than growth in other individual SVs, the total energy was dominated by the much larger set of non-leading SVs from initial to final time. This behaviour is in accord with the previous results of Errico *et al.,* (2001) and Snyder and Hakim (2005) who used more than ten SVs in their analysis of four different synoptic cases, and cyclogenetic perturbations in the context of the quasi-geostrophic Eady model, respectively.

Despite the fact that the total energy was dominated by the slow-developing non-leading SVs, the growth itself was explained in a large proportion by the most rapid SV, which accounted in average for over 50% of the CRCM-perturbations growth in 36 hours.

These results suggest that the growth of CRCM IV was due to the growth of unstable perturbations, the most part being represented by the growth of the first leading SV from the linear operator. The results showed a high similarity between the CRCM perturbations and the first SV after 24- to 36-hour tangent-linear model integration. This SV had initially a predominantly low- to mid-level, westward tilted structure, which was followed by a vertical alignment and amplitude amplification.

The success of the SV analysis is probably due to the fact that, unlike the IV in an ensemble of Global Climate Model simulations, the RCM IV is subject to lateral boundary conditions constraint. As a consequence, the differences between the members in an RCM ensemble behave most of the time as small perturbations with rapid growth and therefore can be decomposed in terms of SVs.

In conclusion, we found that final-time SVs are useful to explain growth of IV in regional climate simulations. It was demonstrated that the projection on the first SV at final time is greater than the projection on other non-leading SVs and there is a high similarity between the CRCM perturbations and the first SV after 24- to 36-hour tangent-linear model integration. However, the initial perturbations do not project well on the initial SVs sustaining the idea that the SVs are outside the attractor of the system at initial time, pointing to areas where solutions do not naturally occur.

We are aware that the results presented above are based on a single winter case and that other cases of large IV should be investigated using the same methodology to confirm our conclusions. Also, other seasons should be considered, such as the summer, although a TLM with a more complete physics – i.e. including simplified parametrizations of moist processes – would probably be necessary, as well as a norm that takes into account the humidity perturbations.

One technical limitation of this study was that we had to use a global model to compute the SVs. To focus on the region of interest, the CRCM domain, we simply restricted the domain of the final-time norm. Ideally, we would have used an SV calculation based on a limited-area model, which would take into account effects of the driving boundary conditions on the growing disturbances of the CRCM.

Acknowledgements

This work is part of the PhD thesis of the first author in Environmental Sciences at Université du Québec at Montréal and supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Graduate Scholarship (CGSD3), as well as by the CRCMD Network supported by the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS) and the Ouranos Consortium. Ouranos and the ESCER Centre provided the computing resources.

We thank Dr. Eugenia Kalnay for her comments on a previous version of the paper and the reviewer for his/her comments that helped improve the manuscript. The first author would like to thank many people for providing the technical support, in particular Michel Giguère, Georges Huard, Nadjet Labassi, Mourad Labassi and Katja Winger.

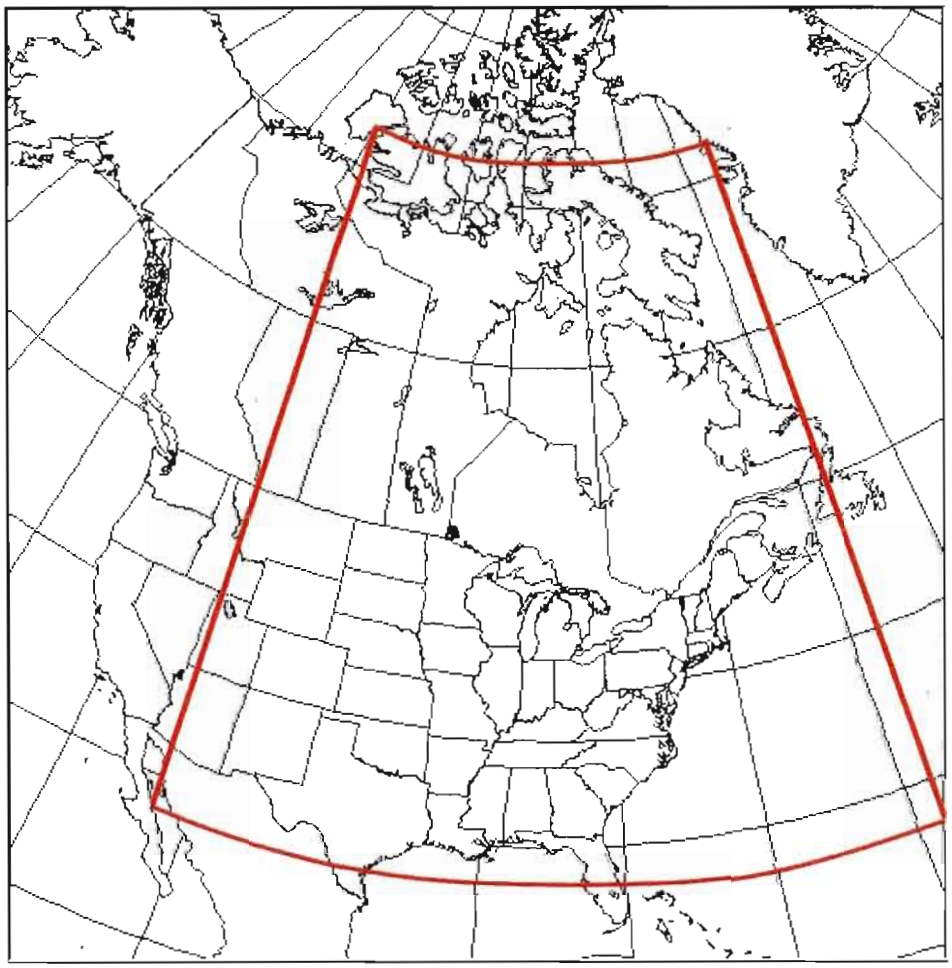


Figure 2.1: CRCM integration domain. The red line shows the initial- and final-norm domain.

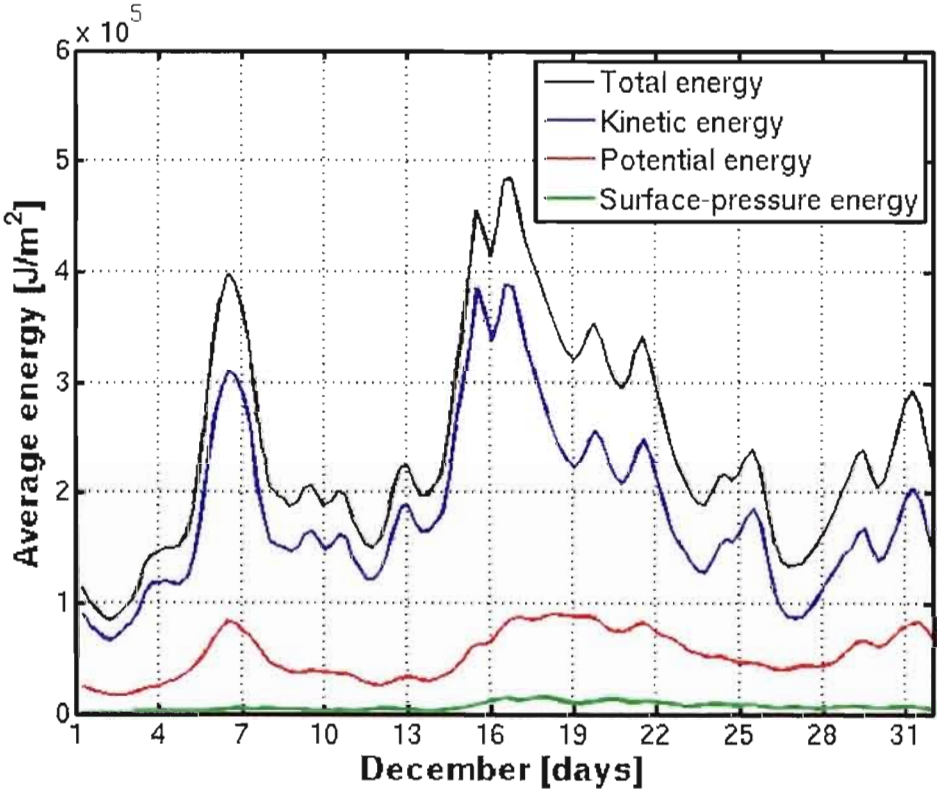


Figure 2.2: Time evolution over the month of December 1992 for CRCM (*black line*) total average energy and its components: (*blue line*) kinetic, (*red line*) potential and (*green line*) surface-pressure energy.

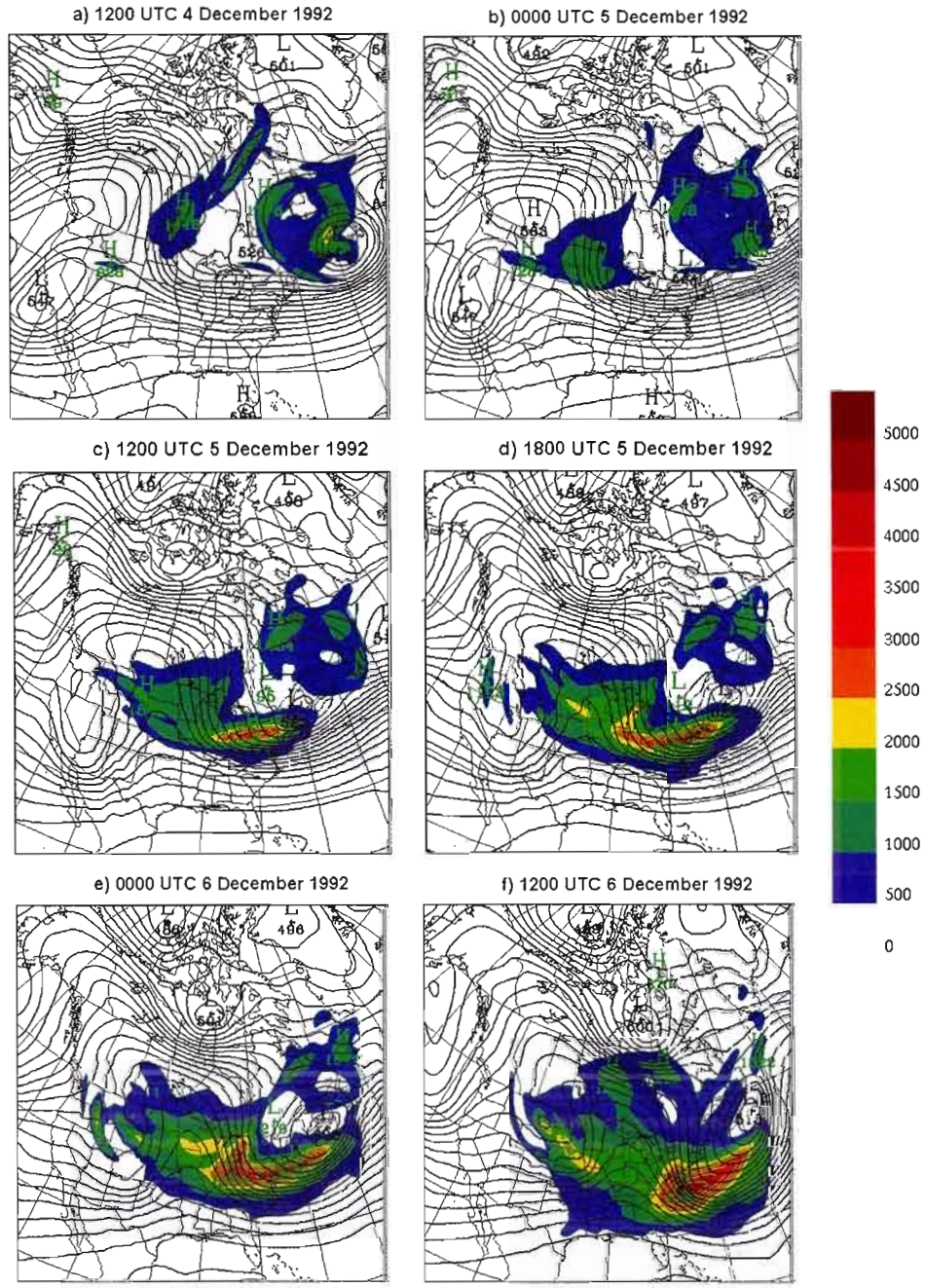


Figure 2.3: 500 hPa geopotential field (*in dam*) for the reference simulation and the average CRCM-perturbation total energy (*in* kJ/m^2) integrated between 100 hPa and surface. The geopotential field (*black contour*) is contoured at 4 dam intervals, while the CRCM-perturbation total energy is plotted in colors.

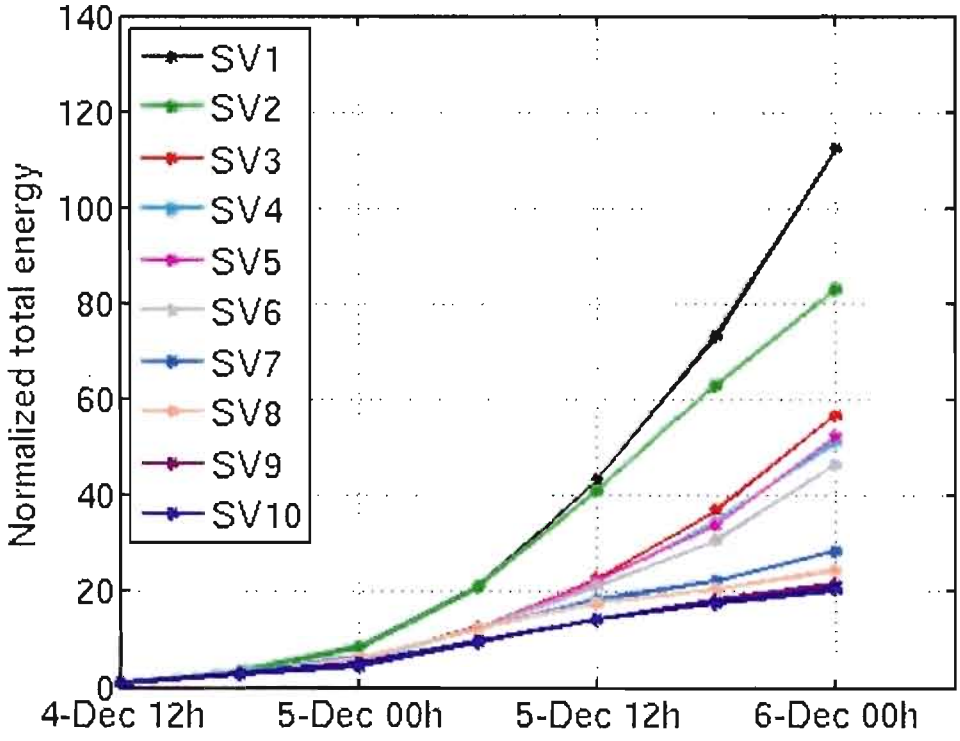


Figure 2.4: Total-energy amplification for the ten SVs during the 36-hours period.

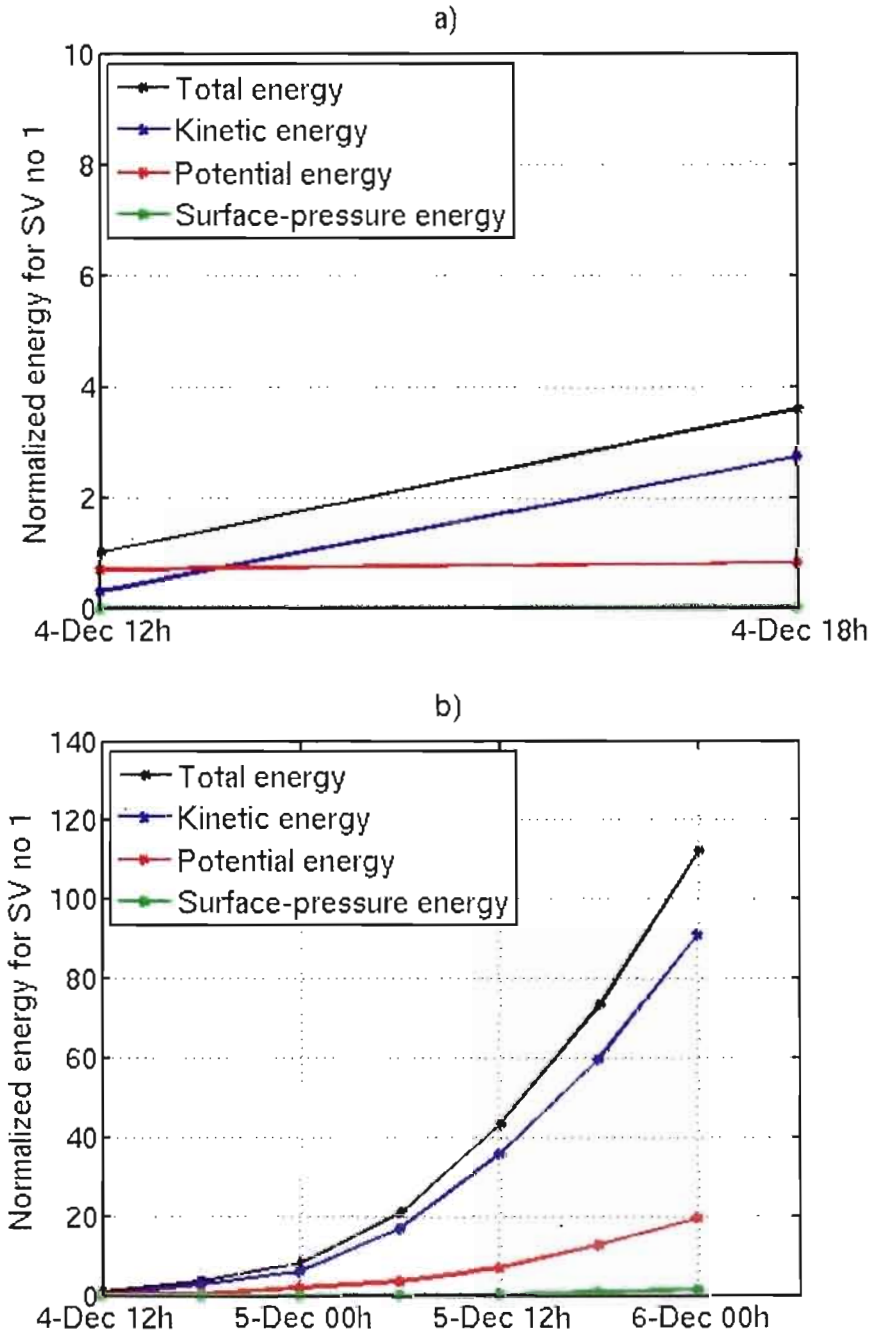


Figure 2.5: 36-hours evolution for the leading SV energy partitioned in kinetic, potential and surface-pressure components. All terms are normalized by the initial total energy. Figure (*a*) shows a zoom on the initial time.

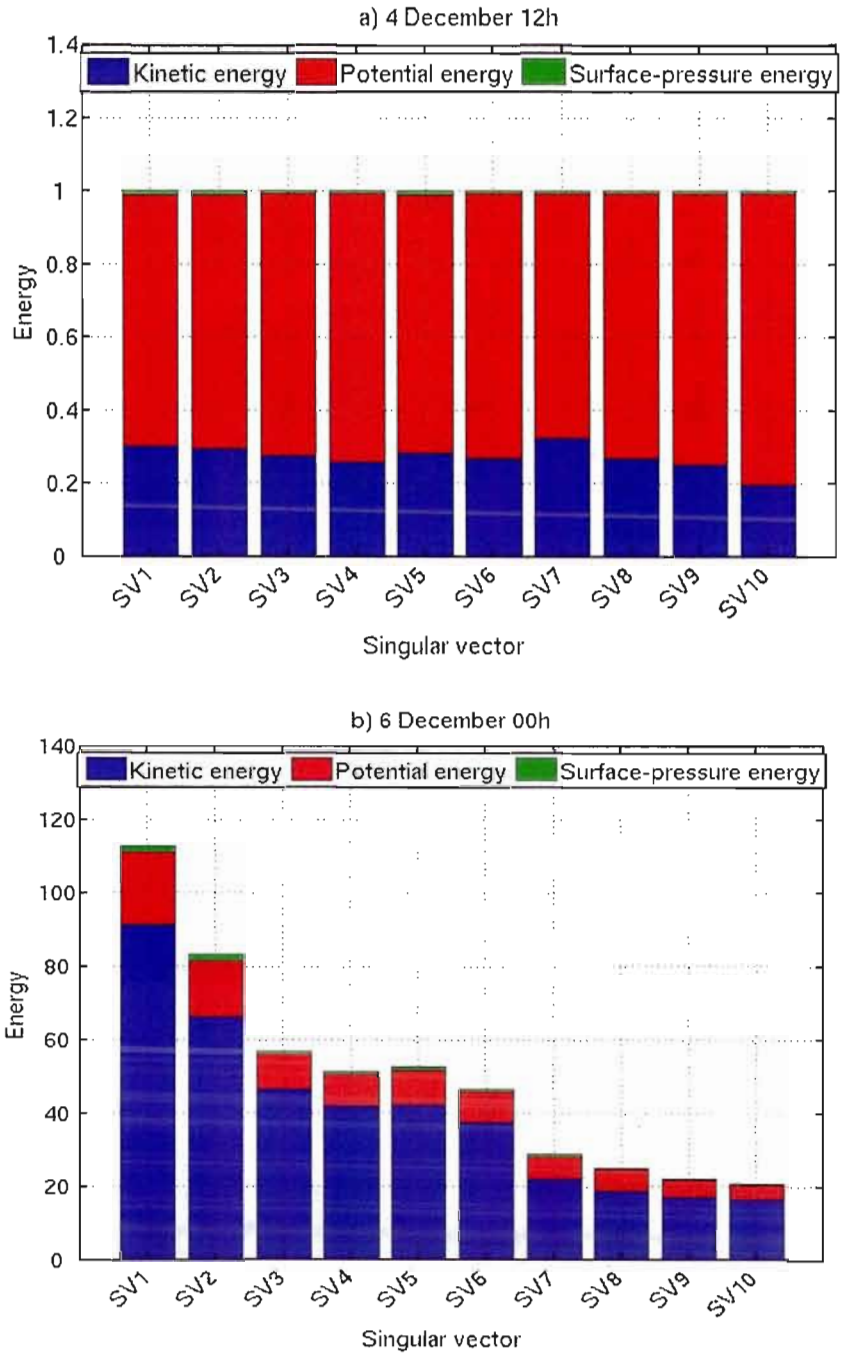


Figure 2.6: SV energy partitioned in kinetic, potential and surface-pressure components at (*a*) initial and (*b*) final time normalized by the initial total energy.

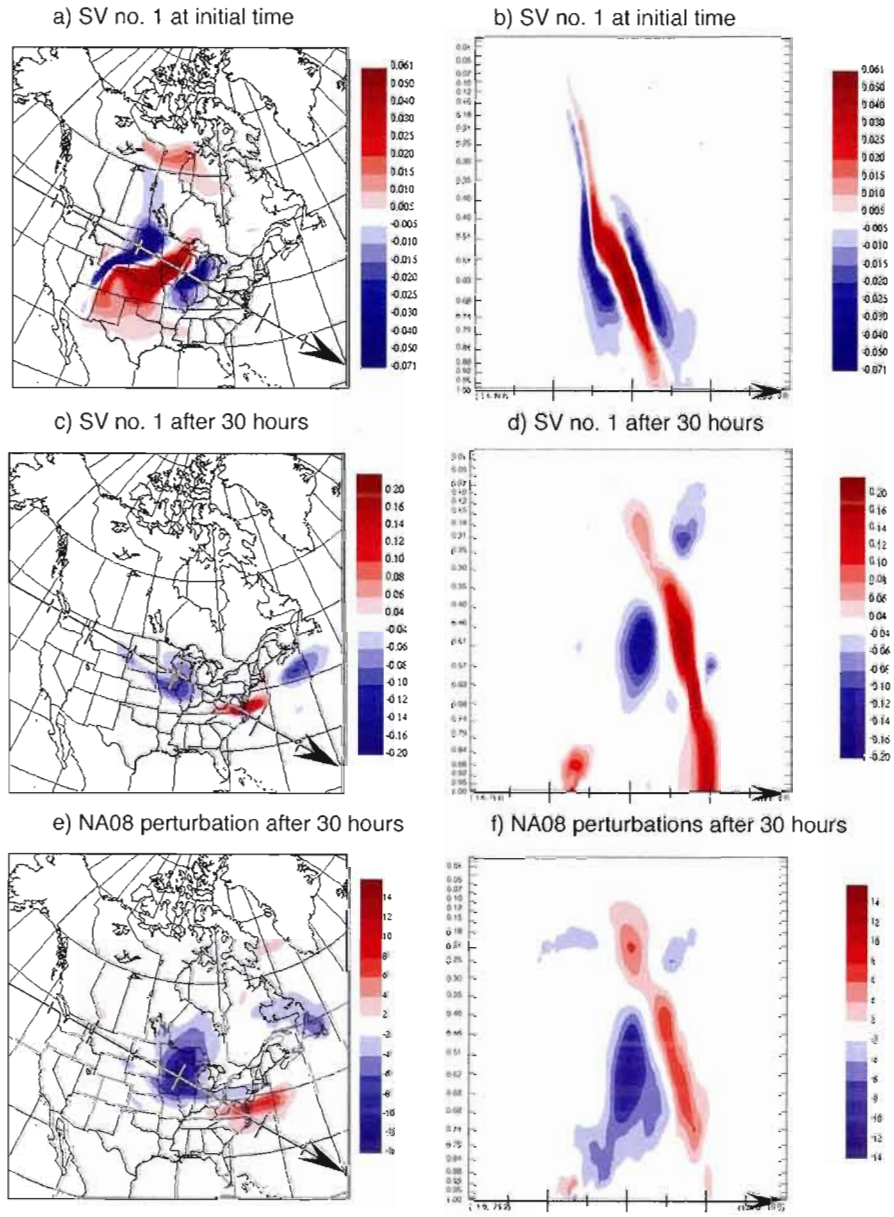


Figure 2.7: Temperature (*in °C*) horizontal structure at 0.688 eta level and vertical cross sections along the arrow corresponding to (*a, b, c, d*) the leading SV normalized by the initial total energy at (*a, b*) initial time and at (*c, d*) after 30 hours and corresponding to (*e, f*) NA08 CRCM perturbation on 1800 UTC 5 December. To facilitate the comparison, the SV was multiplied by -1.

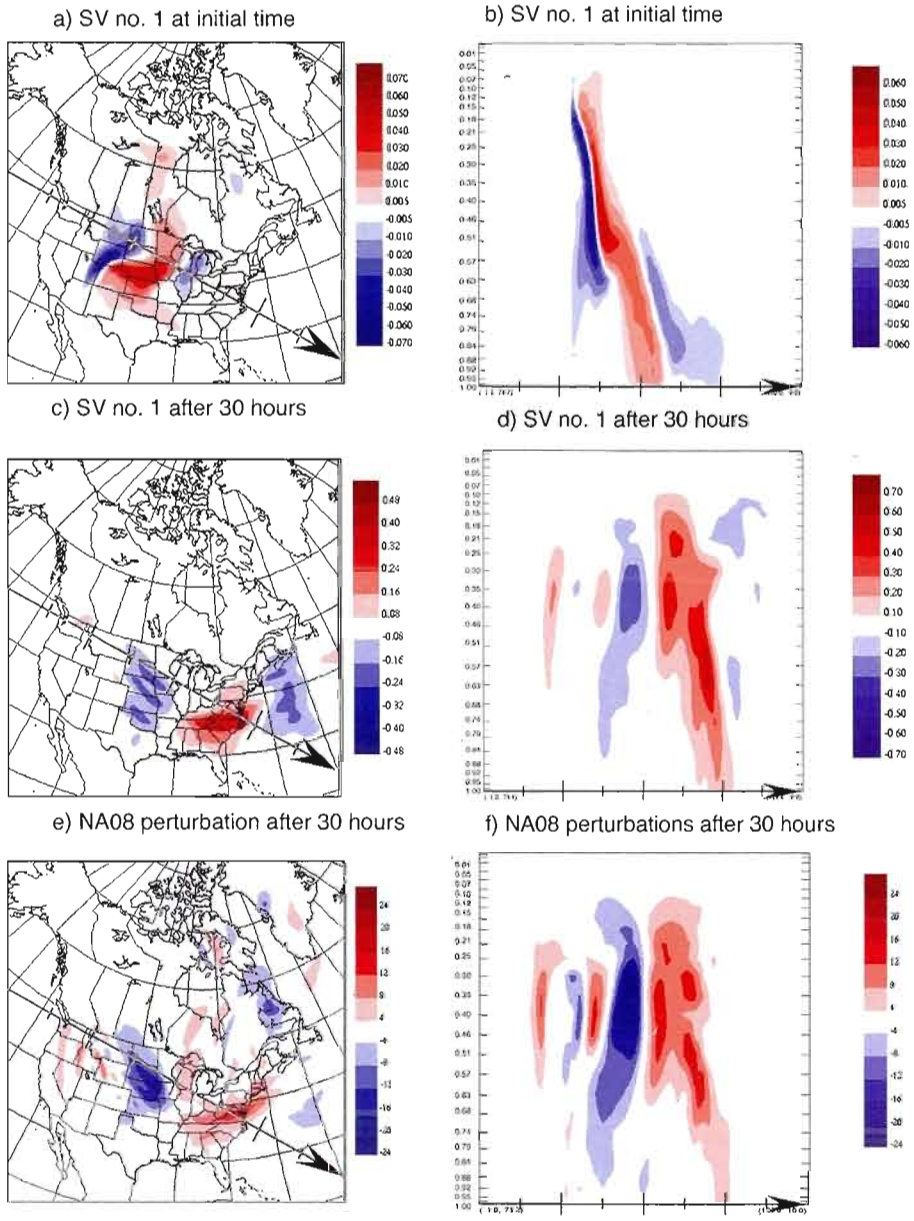


Figure 2.8: Meridional wind *(in m/s)* horizontal structure at 0.688 eta level and vertical cross sections along the arrow corresponding to (*a, b, c, d*) the leading SV normalized by the initial total energy at (*a, b*) initial time and at (*c, d*) after 30 hours and corresponding to (*e, f*) NA08 CRCM perturbation on 1800 UTC 5 December. To facilitate the comparison, the SV was multiplied by -1.

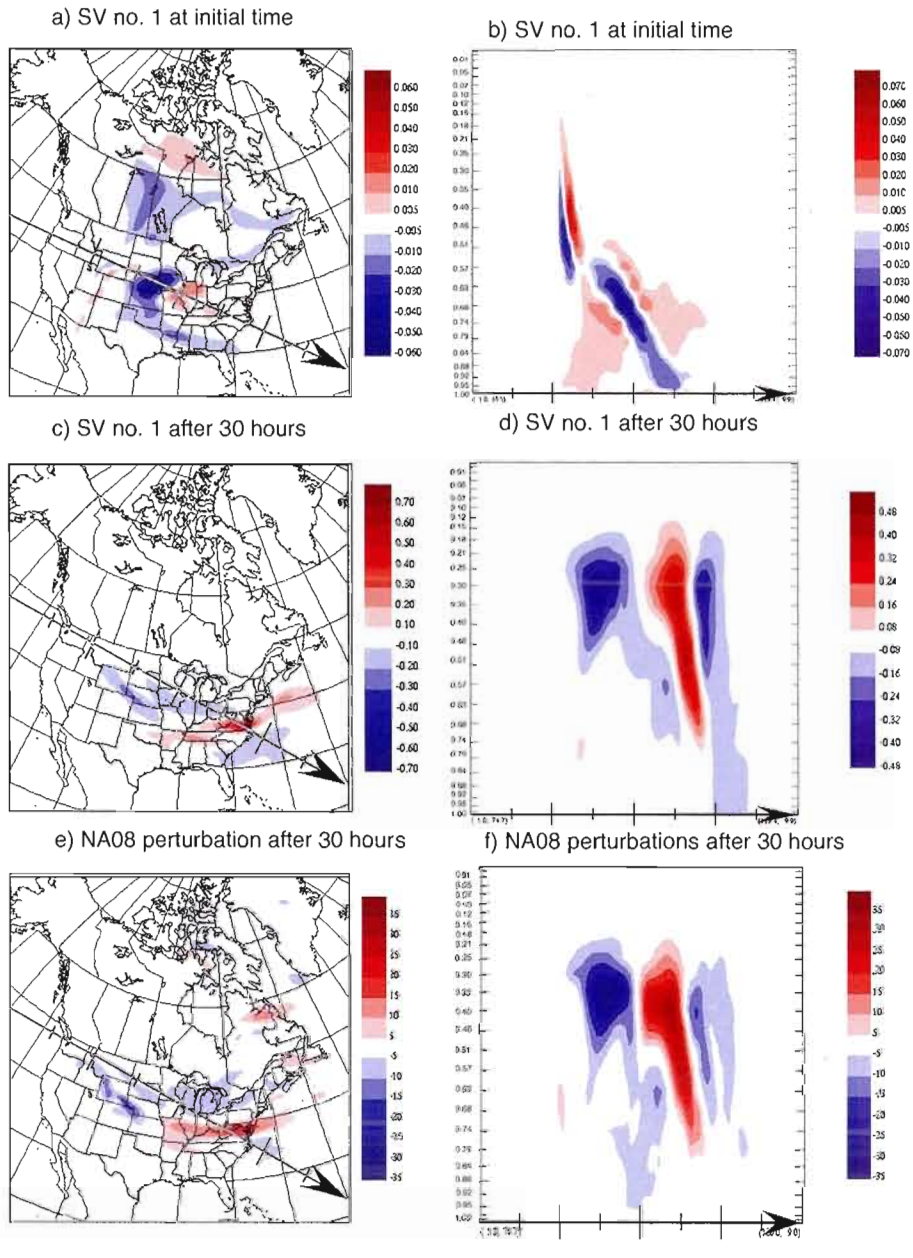


Figure 2.9: Zonal wind *(in m/s)* horizontal structure at 0.688 eta level and vertical cross sections along the arrow corresponding to (*a, b, c, d*) the leading SV normalized by the initial total energy at (*a, b*) initial time and at (*c, d*) after 30 hours and corresponding to (*e, f*) NA08 CRCM perturbation on 1800 UTC 5 December. To facilitate the comparison, the SV was multiplied by -1.

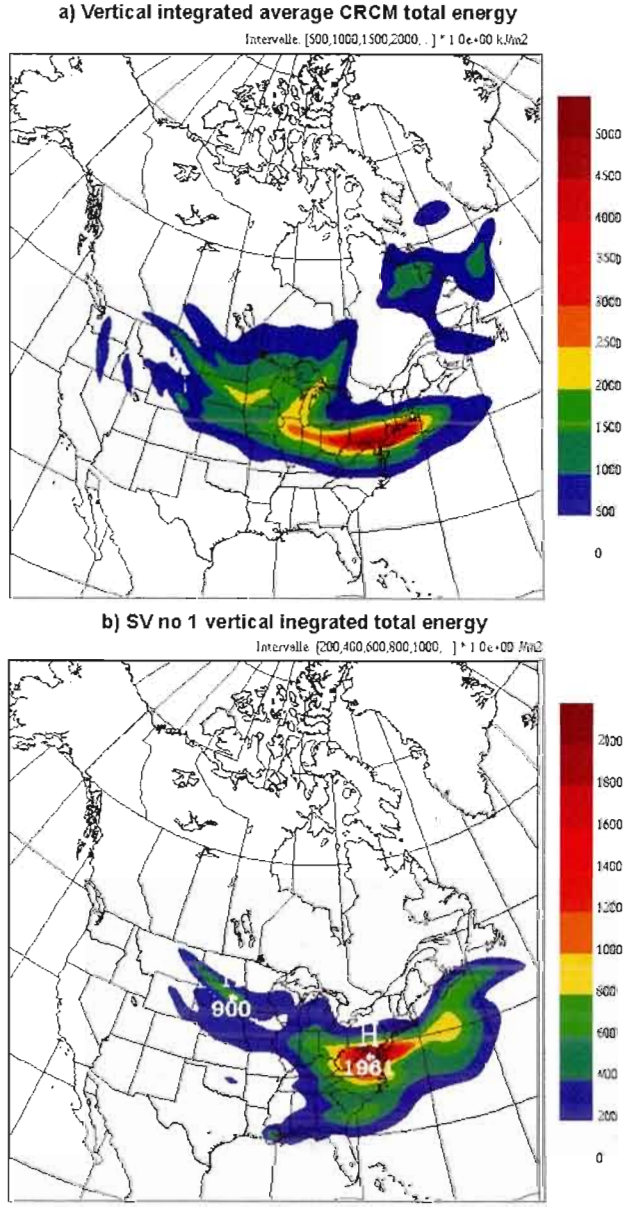


Figure 2.10: Total energy horizontal distribution at 1800 UTC 5 December for (*a*) the CRCM perturbations and (*b*) first SV normalized by the initial total energy.

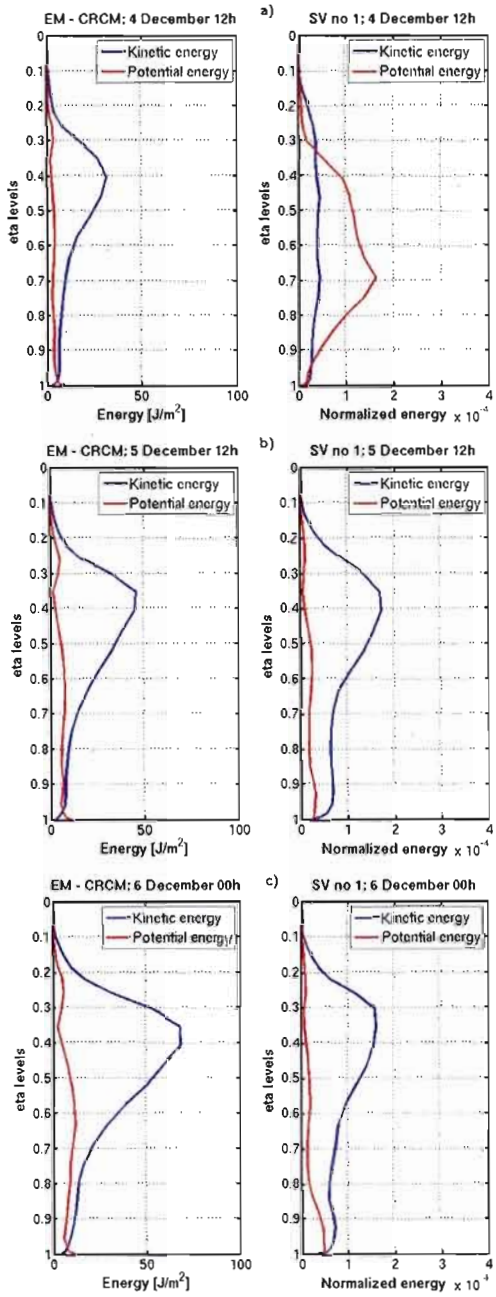


Figure 2.11: Vertical distribution of (*left panels*) first SV kinetic and potential energy and (*right panels*) CRCM-perturbations average kinetic and potential energy at different moments on the 36-h period. Note that SV is normalized by the total energy.

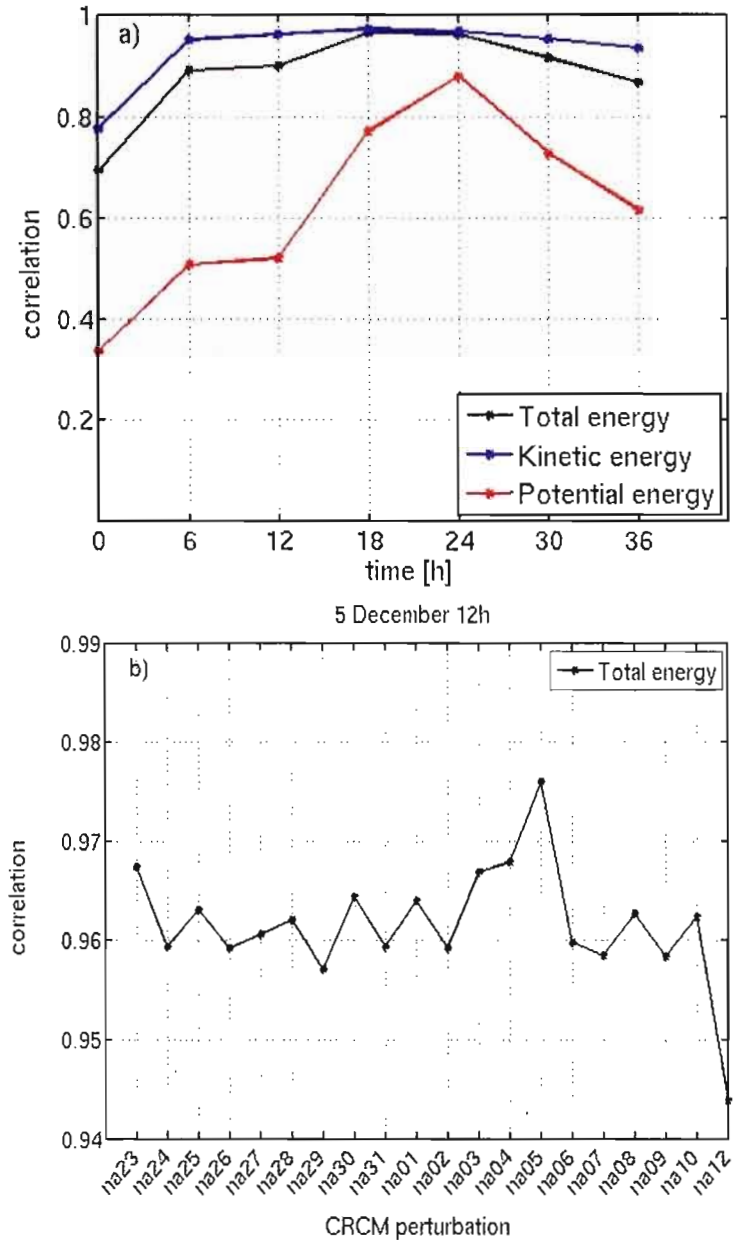


Figure 2.12: (a) Correlation coefficient between the average vertical distributions of CRCM perturbations and first SV energy as function of time for the 36-h period.

(b) Correlation coefficient between the vertical distributions of each CRCM perturbations and first SV total energy by 5 December 12h

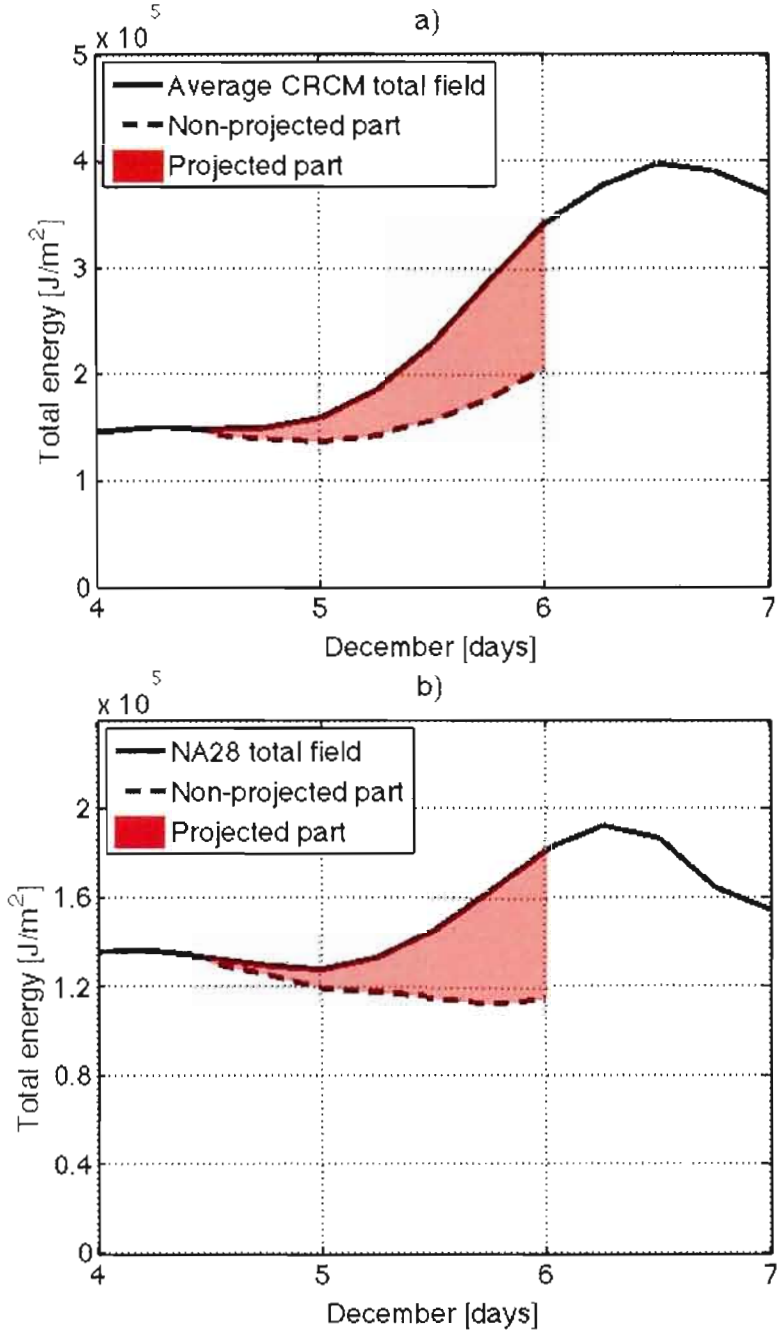


Figure 2.13: (a) The average total energy of CRCM perturbations and (*b*) the total energy of NA28 perturbation: (*dashed black line*) non-projected part, (*red area*) projected part and (*solid black line*) total field.

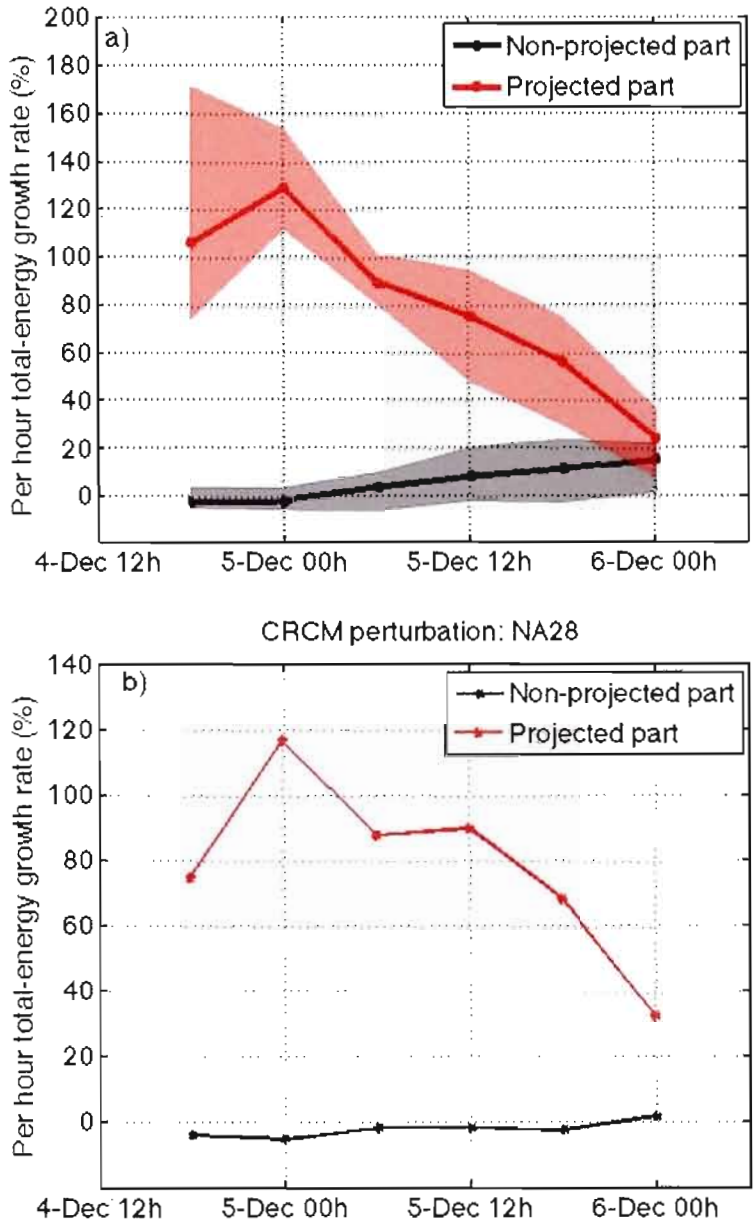


Figure 2.14: 6h-mean growth rate in percentage per hour for (*a*) all CRCM-perturbations and (*b*) NA28 perturbation total-energy (*red line*) projected part and (*black line*) non-projected part.

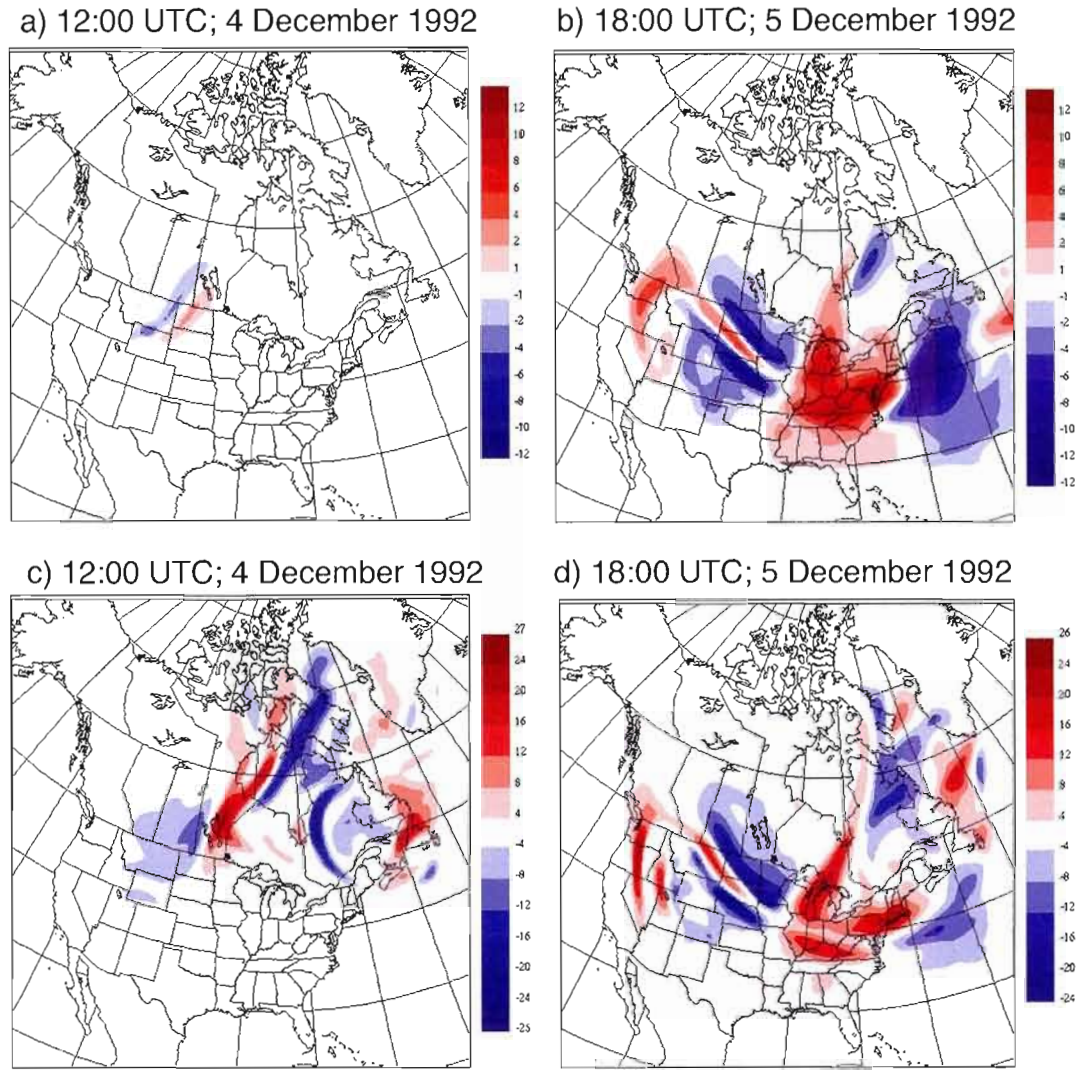


Figure 2.15: Meridional wind (*in m/s*) at approximately 460 hPa for (*a, b*) NA08 projected part on the first SV and (*c, d*) NA08 total field at initial time and after 30 h.

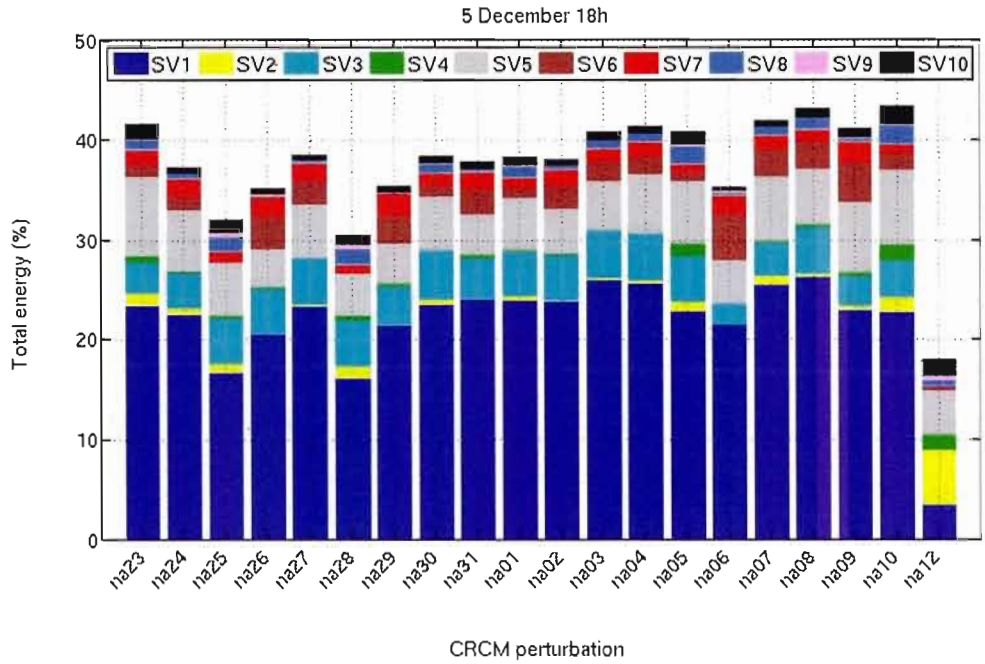


Figure 2.16: Fraction of CRCM-perturbations total energy explained by each SV by 1800 UTC 5 December.

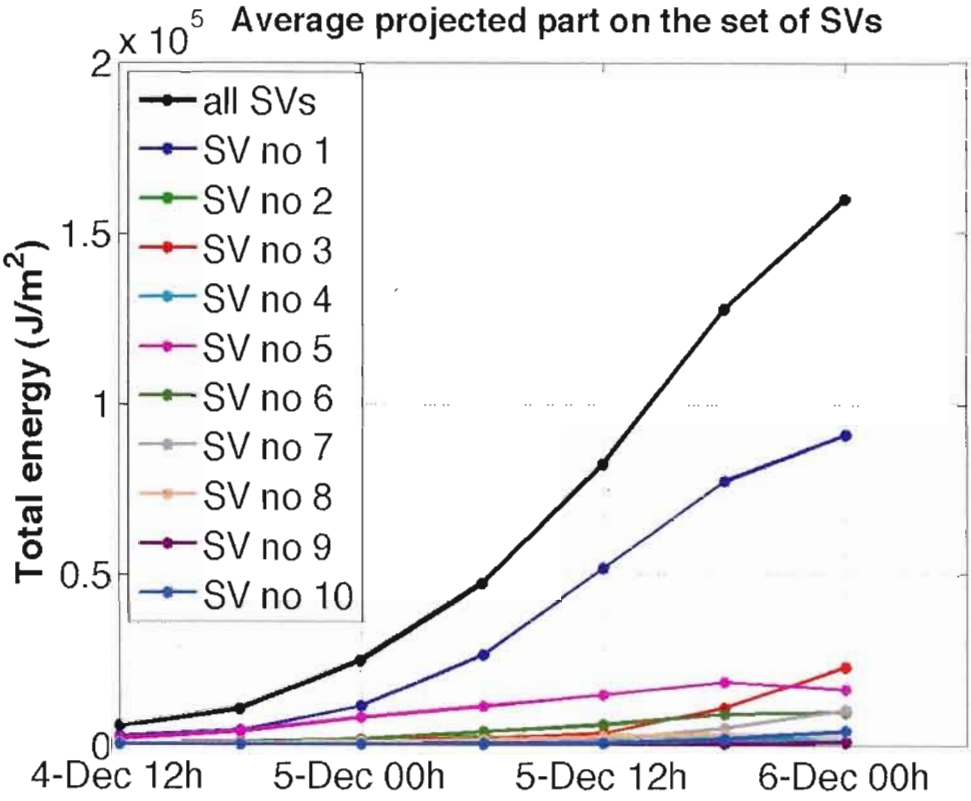


Figure 2.17: Average total-energy temporal evolution for the CRCM-perturbations projected part on each SV.

CHAPITRE III

LES VECTEURS SINGULIERS AVEC LA PLUS FORTE AMPLIFICATION PEUVENT-ILS ESTIMER LES VARIATIONS TEMPORELLES DE LA VARIABILITÉ INTERNÉ D'UN MODÈLE RÉGIONAL DU CLIMAT?

Ce chapitre, rédigé en anglais, est présenté sous la forme d'un article qui sera soumis pour publication dans la revue *International Journal of Climatology*.

L'article poursuit l'analyse comparative de la VI avec les VS avec le plus fort taux d'amplification pour plusieurs périodes caractérisées par différents niveaux de croissance dans la VI. Il remet en question l'utilisation du premier VS, c'est à dire celui avec la plus forte amplification, pour l'estimation des variations temporelles de la VI dans les simulations du MRCC. L'analyse est concentrée sur la variation de la VI et le domaine d'analyse est restreint à la partie centrale du domaine. Cette configuration permet de minimiser l'influence des frontières latérales du modèle sur le calcul. Dans ce contexte, les fluctuations de la VI sont modulées par les processus générateurs de VI ainsi que par le transport de la VI à l'extérieur du domaine d'analyse via le terme de la divergence du flux horizontal par l'écoulement moyen. Le rapport entre la tendance de la VI et la tendance des perturbations MRCC projetées dans les directions des VS indique dans quelle proportion les fluctuations de la VI sont dues à des instabilités représentées par les VS.

CAN THE LEADING SINGULAR VECTORS CALCULATED FROM ONE SIMULATION BE USED TO ESTIMATE THE VARIATIONS OF INTERNAL VARIABILITY IN AN ENSEMBLE OF REGIONAL CLIMATE MODEL SIMULATIONS?

by

Emilia Paula Diaconescu [*,1,3], René Laprise [1,3] and Ayrton Zadra [1,2,3]

[1] *Canadian Network for Regional Climate Modelling and Diagnostics, University of Quebec at Montreal, Montreal, Canada*

[2] *Meteorological Research Division, Environment Canada, Montreal, Canada*

[3] *Centre ESCER (Étude et Simulation du Climat à l'Échelle Régionale)*

To be submitted to

International Journal of Climatology

August 2011

[*] Corresponding author: Emilia Paula Diaconescu
Department of Earth and Atmospheric Sciences
Université du Québec à Montréal (UQAM)
P.O. Box 8888, Succ. Centre-ville
Montréal, Québec
CANADA H3C 3P8

Abstract

In this paper, we question whether the time evolution of singular vectors (SVs) permits to anticipate the time variations of internal variability (IV) in simulations of the Canadian Regional Climate Model (CRCM). We use an ensemble of 21 CRCM simulations and the analysis is focused in the central part of the domain where IV variations are the result of the competition between processes generating IV and the transport of IV out of the analysis domain by basic state flow.

One set with ten SVs and ten sets with five SVs are used to calculate the sources of IV for several 36-hour periods in the first twelve days of December 1992. SVs are computed to maximise the total perturbation energy over an area located within the CRCM domain. For the first five days of December, the IV tendency is dominated by the sources of instability. 36% to 85% of these sources are represented by the four or five most rapid SVs computed with the initial conditions from the CRCM reference state. Only on one instance was the projection dominated by the leading SV. Although there appeared to be sources of instability in the domain, relatively small projections prevail in periods where the IV tendency is dominated by the IV export out of the analysis domain. The fact that the IV in such periods does not project into the directions of the most rapid SVs suggests that the IV growth is due to perturbations that are already in a growing state, and hence, caracterised by a slower growth rate. As a consequence, a large set of SVs would be needed in order to decompose the CRCM sources of IV. For the setup used in this analysis, the leading SV alone does not appear to permit anticipating periods with larger than usual IV growth.

Keywords: internal variability; singular vectors; regional climate model

3.1 Introduction

It is common knowledge that climate models as complex dynamical system are sensitive to small differences in initial conditions. Consider an ensemble of simulations performed with a climate model where everything is the same, except for ICs corresponding to different starting dates. Due to the differences in the ICs, each member of the ensemble will deliver a different solution for a particular future date. The dispersion of model's solutions due to the differences in the ICs is named model internal variability (IV) and can be quantified by the ensemble's inter-member variance.

In the case of an atmospheric global climate model (AGCM), the IV will grow until it reaches the level of natural, transient-eddy variability. In nested regional climate models (RCMs), however, the amplitude of IV is much more reduced compared to an AGCM because the lateral boundary conditions exert some control on the perturbations' growth and will thus limit the IV (e.g. Alexandru *et al.*, 2007; Lucas-Picher *et al.*, 2008a; Laprise *et al.*, 2008; Rapaić *et al.*, 2010). Numerous studies have showed that the IV in RCMs is a function of synoptic conditions: large IV growth usually occur where and when the ensemble-mean state is cyclonic and with significant weather, while calm synoptic periods are characterized by a smaller divergence of the solutions.

Hence RCM IV appears to be the result of a competition between two processes. On one hand, perturbations are excited by the instabilities of the flow. On the other hand, model lateral boundaries (MLB) exert forcing on the simulations toward a unique solution; the strength of this forcing is function of the mean flow speed through the domain, which is strongly dependent on location, season, and domain size (Lucas-Picher *et al.*, 2008a, 2008b). As a result of this competition, the IV in RCM simulations is characterized by episodic growth and decay. The dependence of IV on the domain's size and location, the atmospheric variable under study and on the season has been emphasized in various studies such as those by Rinke and Dethloff (2000), Rinke *et al.* (2004), Alexandru *et al.* (2007), Lucas-Picher *et al.* (2008a, 2008b) and Rapaić *et al.* (2010). The dependence of RCM IV on the instability of the flow was explored in Nikiema and Laprise (2010), who established prognostic budget

equations of the IV of the Canadian RCM for the summer of 1993 over a domain centred on the East Coast of North America. For that particular summer case, the most important contribution to the potential temperature IV growth came from the covariance of potential temperature and diabatic heating from convection and condensation. For the relative vorticity IV tendency the most important terms were associated with horizontal flow components.

Singular vectors (SVs) capture the dynamically most unstable perturbations and finite-time instabilities in the context of a linear model named the tangent linear model (TLM; e.g. Buizza, 1994). By definition, SVs are the orthogonal set of perturbations that, according to the linear theory, will grow fastest over a finite-time interval with respect to some norms specified at initial and final times. Any arbitrary perturbation can be expressed in terms of SVs by projecting it into the directions of the orthonormal set of SVs (Ehrendorfer and Errico, 1995). The choice of norms must be related to the SV application because various choices lead to different initial structures and evolutions (e. g. Buizza, 1998; Palmer *et al.*, 1998; Buehner and Zadra, 2006). The most commonly used norm is the total energy norm. The interval of time over which the growth is calculated is named the optimization time interval (OTI); it varies usually between 24 hours and 48 hours. The choice of OTI is conditioned by the validity of the tangent-linear approximation, which is function of scale (Reynolds and Rosmond, 2003). The small scales are dominated by nonlinear perturbations growth after 12 hours while the synoptic scales remain linear out to 48 hours. Generally, the extratropical total-energy SVs have synoptic scales at final time and are calculated with an OTI of 48 hours.

Usually, the extratropical SVs computed with the total-energy norm have at initial time a westward (upstream) tilt with height and a meridional (from south-west to north-east) phase tilt that diminishes with time (e.g. Buizza, 1994; Montani and Thorpe, 2002; Coutinho *et al.*, 2004). The vertical westward tilt indicates baroclinic conversion of available potential energy into kinetic energy, while the meridional tilt against the horizontal shear denotes a barotropic mechanism for SV growth characterized by the transformation of basic-state kinetic energy into the perturbation kinetic energy.

In a previous study (Diaconescu *et al.*, 2011), we expressed the IV in Canadian RCM (CRCM) simulations in terms of perturbations from a reference state and we compared the

growth of IV for a period starting on 4 December at 12:00 UTC until 6 December at 00:00 UTC with a set of SVs computed using the Canadian Global Environmental Multiscale (GEM) TLM with total-energy norms at initial and final times confined to the CRCM domain. We have found that the first ten evolved SVs account on average for 70% of the CRCM IV growth in 36 hours. That study focused on a period with rapid growth in IV.

The present study goes beyond, by questioning whether the SVs can explain the IV fluctuations over time. The study also investigates whether the leading SV (that is the SV with largest growth rate) can be used as a predictor of IV growth. The tendency equation for the perturbation total energy is used. The methodology and a brief overview on the CRCM ensemble and the construction of the SVs are presented in section 3.2. The results of this investigation are shown in section 3.3, while section 3.4 summarizes the main findings and provides the conclusions.

3.2 Methodology and models descriptions

3.2.1 The CRCM ensemble and the IV tendency

The issue of model IV is addressed by constructing an ensemble of simulations with different initial conditions. Because at climate scales, the model's response to initial atmospheric perturbations is independent of the magnitude as well as the type of the perturbation, the ensemble is usually constructed by offsetting the initialization time of each simulation by 6 or 24 hours.

For this study, an ensemble of twenty-one integrations, with an offset of 24 hours in initial conditions, were realized with version 5 of the Canadian Regional Climate Model (CRCM5), developed by the Canadian Regional Climate Modelling and Diagnostics Network, in collaboration with the Meteorological Service of Canada (MSC) (Zadra *et al.*, 2008). CRCM5 is a limited-area nested version of the Global Environmental Multiscale (GEM) model employed for numerical weather prediction at the MSC (Côté *et al.*, 1998; Yeh *et al.*, 2002). CRCM5 operates on a rotated latitude-longitude mesh. In this study, CRCM5 has 53 nonuniformly spaced hybrid levels in the vertical, on a 120 x 120 grid-point mesh with spacing of 0.5 degree, and a time step of thirty minutes. The integration area is centred on the

North American continent and can be visualized in Figures 3.1 and 3.3. A ten grid-point wide Davies-type lateral boundary relaxation zone (Davies 1976) was applied. The ten-points buffer zone is indicated in Figures 3.1 and 3.3 by the green line. No large-scale spectral nudging (Alexandru *et al.,* 2009) has been applied in this study.

The initial and lateral boundary conditions have been taken from the ECMWF (European Centre for Medium-range Weather Forecasts) reanalysis project ERA40 (Uppala *et al.,* 2005) available every six hours. The ocean surface conditions were prescribed. The sea surface temperatures (SST) and sea-ice were interpolated from the Atmosphere Model Intercomparison Project (AMIP2; Gleckler, 1996), available on a one-degree latitude-longitude grid for monthly mean values.

In our configuration, the model uses the Kain and Fritsch (1990) deep convection scheme, the Kuo (1965) transient scheme for shallow convection (Bélair *et al.,* 2005), the Sundqvist scheme for large-scale condensation (Sundqvist *et al.,* 1989), and the Interactions Soil-Biosphere-Atmosphere (ISBA) as the land-surface scheme (Bélair *et al.,* 2003). The radiation package for solar and terrestrial radiation is based on the correlated-K approach (Li and Barker, 2005). Subgrid-scale orographic gravity-wave drag is due to McFarlane (1987) and low-level orographic blocking is described in Zadra *et al.* (2003).

The 21 member simulations employed the same model configuration, except for the ICs that correspond to different starting dates between October 23 to November 12 at 00:00 UTC; the corresponding runs are noted as *NA23*, *NA24*, ... , *NA31*, *NA01*, ... , *NA11*, *NA12*, the number representing the ICs' days. Model output was archived every six hours and the analysis is restricted to the first two weeks of December 1992. Because the simulations use the same driving fields and the same surface fields in the ICs, which are interactive afterwards, there are no lateral and lower boundary conditions perturbations. By design, the dispersion of model simulations is due only to the differences in the atmospheric initial conditions.

The IV is defined as the spread between the members in an ensemble with respect to a reference state:

$$\sigma_{IV}^2(x,y,z,t)=\frac{1}{M-1}\sum_{m=1}^{M}\left[x'_m(x,y,z,t)\right]^2 \tag{3.1}$$

where M corresponds to the total number of members in the ensemble of simulations, $x'_m(x,y,z,t)$ represents the perturbation of simulation number m (X_m) with respect to a reference state (X_{ref}):

$$x'_m = X_m - X_{ref}. \tag{3.2}$$

The ensemble mean is often used as the reference state to evaluate the IV (e.g., Alexandru *et al.*, 2007; Šeparović *et al.*, 2008; Rapaić *et al.*, 2010). But because the ensemble mean is not generally a solution of the field equations, we opted in this study to use as reference a specific member of the ensemble, and this reference state was used to compute the SVs. The member *NA11* was chosen as reference simulation because it was the closest one to the ERA40 driving data.

We mention that, in the case of an AGCM, the spread of an ensemble is very large, approaching the value of natural transient-eddy variability. As a concequence, the members will be very different and can have different stability characteristics. However, in the case of an RCM, the spread of an ensemble is limitated by the model lateral boundaries, which forces all members to have the same solution as the driving field in the region of lateral boundaries. Therefore, the RCM members do not have time to diverge very much and will present similar features. The differences between the members of the RCM ensemble can be regarded in this case as perturbations from a time-evolving reference state.

In the following the ensemble will be treated in terms of perturbation from the NA11 simulation and the prime will be dropped to simplify notation. The perturbation total dry energy can be computed as

$$E_m = \iiint_{Vol} \rho\,(u_m{}^2 + v_m{}^2)dVol + \iiint_{Vol} \rho\,(\frac{c_p}{T_r}T_m{}^2)dVol + \iint_{A}(\frac{RT_r}{p_r g}p_{sm}^2)dA. \tag{3.3}$$

where (u_m, v_m, T_m, p_{sm}) represent m-th member perturbation horizontal wind, temperature and surface pressure fields, A is the model horizontal domain, Vol is the model volume, $g = 9.806$ m s^{-2} is the gravitational acceleration, $R = 287.04$ J K^{-1} kg^{-1} is the gas constant for

dry air, c_p = 1005.46 J K^{-1} kg^{-1} is the specific heat for dry air at constant pressure, T_r = 300 K is a reference temperature and p_r = 1000 hPa is a reference surface pressure. The three terms on the right-hand side represent the kinetic, potential and surface-pressure components of the total perturbation energy, respectively.

The ensemble-mean perturbation energy defines IV here:

$$E_n = \frac{1}{M-1} \sum_{\substack{m=1 \\ m \neq n}}^{M} E_m \tag{3.4}$$

where n refers to the reference simulation (*NA11* in our case).

The time evolution for the vertically and horizontally integrated perturbation energy (Figure 3.2) shows that IV is characterized by episodes of growth and decay, as noted previously for various fields such as temperature, geopotential or precipitations (e.g. Alexandru *et al.*, 2007; Nikiema and Laprise, 2010). In these simulations, the first event with important IV growth begins on 4 December at 12:00 UTC and reaches a maximum on 6 December at 12:00 UTC. In the following week, the IV generally decreases and there are only three short periods with some IV growth. The next event with important growth begins on 13 December at 12:00 UTC. In all cases, the surface-pressure component of energy is very small compared to the kinetic and potential terms, so from now on, the perturbation energy will include only the kinetic and potential contributions.

The horizontal distribution of the vertically integrated ensemble-mean perturbation energy is shown in Figure 3.1 for 5, 6 and 7 December, all at 06:00 UTC, together with the reference state 500-hPa geopotential field. The figure reveals that IV grows while it is transported by the basic state flow through the domain. In their study, Lucas-Picher *et al.* (2008b) showed that the magnitude of IV was well correlated with the time that an air parcel spends inside the CRCM domain, a longer residence time favouring larger IV. Figure 3.1c reveals the transport of the IV maximum toward the eastern boundary where it is forced to zero because all members share the same lateral boundary conditions.

Similar feature was also highlighted by the work of Nikiema and Laprise (2010). Considering that the basic state acts as a reservoir of energy, the IV variations can be

explained as follows. The IV is arising through rapidly growing perturbations developed in dynamically unstable regions; they are further transported by the reference state, and will finally be forced to disappear when they approach the MLB. Following this idea we can write an equation describing the perturbations evolution inside the RCM domain as:

$$\frac{dE}{dt} = S_{\text{int}} + S_{MLB}\,. \tag{3.5}$$

where E represents the perturbations energy. The left-hand side term of the equation represents the time variation of perturbation energy following the flow, and the right-hand side terms represent the sources and sinks of perturbation energy that are separated into the sinks of IV caused by the MLB forcing (S_{MLB}), and the sources and sinks within the free domain (S_{int}). The equation can be written using the local variation form:

$$\frac{\partial E}{\partial t} + \mathbf{V}\cdot\nabla E + \omega\frac{\partial E}{\partial p} = S_{\text{int}} + S_{MLB}\,. \tag{3.6}$$

where $\mathbf{V} = (u, v)$ is the basic-state horizontal wind vector. The first term on the left-hand side represents the local variation of perturbation energy and the second and third terms gives the horizontal and vertical advection of energy. By using the continuity equation,

$$\nabla\cdot\mathbf{V} + \frac{\partial\omega}{\partial p} = 0, \tag{3.7}$$

the Eq. (3.6) can be transformed in flux form as

$$\frac{\partial E}{\partial t} + \nabla\cdot(\mathbf{V}E) + \frac{\partial}{\partial p}(\omega E) = S_{\text{int}} + S_{MLB} \tag{3.8}$$

where $\mathbf{V}E$ and ωE are the horizontal and vertical fluxes of perturbation energy.

By vertically integrating the equation (3.8), assuming ω to be zero at surface and at the top of the model, the equation becomes:

$$\frac{\partial E_h}{\partial t} = (S_{\text{int}} + S_{MLB})_h - \nabla\cdot(\mathbf{V}E)_h \tag{3.9}$$

where the indices "h" indicates a quantity vertically integrated. The equation shows that the local variation of the perturbation energy is given by the algebraic sum of sources and sinks

of energy and vertically averaged horizontal transport by the mean flow.

Consider now the integral over the entire horizontal domain (A) of the model:

$$\iint_A \frac{\partial E_h}{\partial t}\, dA = \iint_{dA} (S_{\text{int}} + S_{MLB} - \nabla \cdot (\mathbf{V}E))_h \, dA \qquad (3.10)$$

If one were to integrate this equation over the entire model domain, the divergence term would be zero because E=0 at the lateral boundaries of the model, and the perturbations variation would be given only by the sources and sinks of energy:

$$\iint_A \left(\frac{\partial E_h}{\partial t} \right) dA = \iint_{dA} (S_{\text{int}} + S_{MLB})_h \, dA \qquad (3.11)$$

However it is difficult to account for the sinks of IV caused by the MLB forcing. So we chose a different approach, restraining the analysis over a "free domain" (A_{free}), i.e. a region excluding the MLB. Under these conditions, the A_{free}-domain integral becomes:

$$\iint_{A_{free}} \left(\frac{dE_h}{dt} \right) dA_{free} = \iint_{A_{free}} \left(\frac{\partial E_h}{\partial t} \right) dA_{free} + \iint_{A_{free}} \left(\nabla \cdot (\mathbf{V}E)_h \right) dA_{free} = \iint_{A_{free}} (S_{\text{int}})_h \, dA_{free} \,. \quad (3.12)$$

Over A_{free}, the sinks caused by MLB forcing are zero but the integration involves flux divergence, corresponding to the horizontal export of IV out of the domain by the mean flow.

Figure 3.3a displays the horizontal distribution of the vertically integrated perturbation total-energy divergence term ($\nabla \cdot (\mathbf{V}E)_h$) on 6 December 1992 at 06:00 UTC, while Figure 3.3b shows the perturbation total-energy local tendency for the next six hours ($\frac{\partial E_h}{\partial t}$). The green line excludes the buffer zone of ten points, while the magenta line shows the region of 96 x 96 horizontal grid points that we have chosen as "free domain" away from the MLB forcing. The divergence term highlights the transport of perturbation energy into the east buffer zone; however, the tendency term does not present any growth into this region. This is because dissipation processes there cancel the convergence into the region of MLB. Similar behaviour has been noted at other times in the simulation.

Figure 3.4 shows the two-week time evolution of the local $\iint_{A_{free}} \left(\frac{\partial E_h}{\partial t} \right) dA_{free}$ (red curve)

and total (black curve) tendencies and the convergence term $\iint_{A_{free}} (-\nabla \cdot (\mathbf{V}E)) dA_{free}$ (blue curve), all horizontally integrated over the analysis domain. The IV total variation was calculated by summing the local tendency and divergence terms, as in Eq. 3.12. All terms are expressed in J $m^{-2}h^{-1}$ and are computed at each six hours, the CRCM archival time. The local tendency is computed as the change over the next six hours and the divergence term is represented as the average value over the six-hour time interval.

According to (3.12), the total variation is equal to the sum of the sources and sinks within the free domain. Positive values indicate that the sources predominate over the sinks during the period. On the other hand the local tendency varies in time between positive and negative values. Rapid decrease of IV correspond to export of IV out of the analysis region, while periods with increase in IV are dominated by the IV sources.

3.2.2 Construction of SVs

The leading SVs represent the fastest growing perturbations from a given trajectory in parameter space, in a linear sense (e.g., Palmer *et al.*, 1998; Buehner and Zadra, 2006). Consider a nonlinear model **H** and a state vector **X**(t). Running the model from time t_0 to t will generate a trajectory from an initial state $\mathbf{X}(t_0)$ to **X**(t):

$$\mathbf{X}(t) = \mathbf{H}(\mathbf{X}(t_0)). \tag{3.13}$$

Let $x(t)$ be a small perturbation from the non-linear model trajectory **X**(t). For small perturbations, the model equations can be linearized and the perturbation evolution can be computed by using the forward-tangent linear model (TLM):

$$x(t) = \mathbf{L}(t_0, t)x(t_0). \tag{3.14}$$

where $\mathbf{L}(t_0, t) = \left.\frac{\partial \mathbf{H}}{\partial \mathbf{X}}\right|_{\mathbf{X}(t)}$. The magnitude of perturbations can be scaled in terms of a norm. Consider $\mathbf{E}_t$ as the perturbation norm at time t:

$$\|x(t)\|_{E_t}^2 = \langle x(t); E_t x(t) \rangle. \tag{3.15}$$

The ratio of norms at final and initial times defines the amplification factor in the

interval of time *(t_0, t)*:

$$\lambda^2 = \frac{\|x(t)\|^2_{E_t}}{\|x(t_0)\|^2_{E_0}} = \frac{\langle \mathbf{L} * E_t \mathbf{L} x(t_0); x(t_0) \rangle}{\langle x(t_0); E_0 x(t_0) \rangle}. \tag{3.16}$$

where $\mathbf{E_0}$ is the norm at initial time and $\mathbf{L}^*$ is the adjoint of $\mathbf{L}$ with respect to the norm $\mathbf{E}$, and it has the property that $\langle x; \mathbf{L}y \rangle = \langle \mathbf{L} * x; y \rangle$ for all *x, y* in the space. The most rapidly growing perturbations, the leading SVs, are computed to maximize the ratio defined by Eq. (3.16). SVs are the solution of the following generalized eigenvalue problem:

$$(\mathbf{L} * E_t \mathbf{L}) y_i(t_0) = \lambda_i^2 E_0 y_i(t_0), \tag{3.17}$$

The initial SVs ($y_i(t_0)$) form an $\mathbf{E_0}$-orthonormal basis and can be ordered according to their growth potential (λ_i^2), with the fastest growing structure being called the first SV. The evolved or final SVs form an $\mathbf{E_t}$-orthogonal set at optimization time and can be obtained by integrating the TLM.

Because the software required for the calculation of the SVs on a limited-area domain was not yet ready at the time of our investigation, the SVs employed in this study have been computed using the TLM and adjoint versions of global GEM model, originally developed for 4D variational data assimilation system (Gauthier *et al.,* 2007). SVs are computed with a total dry energy norm at both initial and final times:

$$< x, Ex >= \iiint_{Vol} \rho \, (u^2 + v^2) dVol + \iiint_{Vol} \rho \, (\frac{c_p}{T_r} T^2) dVol + \iint_{A} (\frac{RT_r}{p_r g} p_s^{\,2}) dA \tag{3.18}$$

where $x = (u, v, T, p_s)$ represents a perturbation of the model state vector comprising horizontal wind components, temperature and surface pressure perturbations. The vertical integral is restricted to the levels situated above approximately 100 hPa in order to focus on the troposphere (Zadra *et al.,* 2004).

The TLM employs a uniform horizontal resolution of 1°, 28 eta levels and a time step of 30 minutes, and uses only simplified physics consisting of vertical diffusion. A detailed description of the available simplified parameterizations in the GEM TLM and their impact on SVs computed with the total-energy norm can be found in Zadra *et al.* (2004). The

evolution trajectory was generated using the nonlinear full-physics global GEM model, while the ICs were constructed by combining the reference CRCM simulation (*NA11*) over the region encompassed by the CRCM integration, with the ERA40 reanalysis elsewhere on the globe.

Eleven sets of SVs are used in this study. The first set, noted *SetA*, has the following characteristics:

- The ICs correspond to 4 December 1992 at 12:00 UTC.
- The initial norm is limited to a sub-domain within the CRCM integration area (the blue region in Figure 3.5), which permits the selection of the SVs situated within the region of CRCM simulations.
- The number of SVs is limited to the first ten that are archived every six hours.

The other ten sets of SVs are noted *SetB1*, *SetB2, ..., SetB10* and, with the exception of the ICs that correspond to different dates (Table 3.1), they all have the same setup:

- They have a global initial time norm.
- The number of SV is limited to the first five SVs that are archived only at initial and final times.

The number in the name of these ten sets corresponds to the date of ICs (see Table 3.1). All eleven sets of SVs have an OTI of 36 hours and final norms limited to the CRCM sub-domain represented by the blue zone in Figure 3.5. The *SetA* was already used in a previous study to measure the growth of CRCM IV in terms of SVs. Here it is used complementary to the *SetB1*, *SetB2, ..., SetB10* because it counts a larger number of SVs that are archived every six hours, providing supplementary information for a period with rapid growth in the CRCM IV.

All SVs are available in the form of perturbations in horizontal wind components, temperature and surface pressure fields, on grid spacing of 1° and 28 eta levels. In order to facilitate the comparison with the CRCM fields, the SV were further interpolated on the 0.5° CRCM grid and all the results are calculated with respect to the total-energy norm defined over the CRCM analysis domain.

3.3 SVs versus the CRCM IV

Consider a set truncated to only the first N fastest growing SVs. The N-dimensional subspace of the system's phase space spanned by the final-time SVs is:

$$G(t)=\{\hat{y}_j(t), j=1,N\}, \tag{3.19}$$

where $\hat{y}_j(t)$ represents the SVs normalized to have the unity norm. The CRCM perturbations $\{x_k(t), k = 1, M-1\}$ can be compared to the first N SVs by projecting them in the directions of the SVs (e.g. Buizza *et al.*, 1993; Buizza, 1994). Only the fast growing part $\tilde{x}_k(t)$ of the CRCM perturbation $x_k(t)$ will project into this subspace:

$$\tilde{x}_k(t)=\sum_{j=1}^{N}\alpha_{kj}\hat{y}_j(t) \quad , \tag{3.20}$$

with $\alpha_{kj}=<x_k(t),E_t\hat{y}_j(t)>$ representing the projection coefficients. The square of the projection coefficient, α_{kj}^2, gives the amount of energy of the k^{th} CRCM perturbation that is explained by the j^{th} SV. The sum of the squares of the projection coefficients with a fixed index k represents the energy of the projected part of the CRCM perturbation $x_k(t)$,

$$E(\tilde{x}_k)=\left\|\tilde{x}_k(t)\right\|^2=\sum_{j=1}^{N}\alpha^2{}_{kj}, \tag{3.21}$$

while the fraction of the CRCM perturbation explained by the first N SVs is given by the ratio between the norm of the projected part and the norm of the full perturbation:

$$p_k(t)=\frac{\left\|\tilde{x}_k(t)\right\|^2}{\left\|x_k(t)\right\|^2}. \tag{3.22}$$

In this study, each set of SVs is computed to maximise the total energy over an area situated into the CRCM free domain after thirty-six hours of GEM linear integration.

First, we analyse the *SetA* of SVs, which marks the beginning of a period with important growth in the CRCM IV.

3.3.1 SetA of SVs and the IV tendency

As mentioned before, *SetA* of SVs was already used in a previous study (Diaconescu *et al.*, 2011) to verify whether the growth of IV for the period from 4 December at 12UTC till 6 December at 00UTC can be explained by the first ten SVs. It has been found that there is a high similarity between the CRCM perturbations and the leading SV after twenty-four to thirty-six hours of the TLM integration. The first final-time SV accounts on average for 23% of the CRCM IV and for 50% of the projected part. However, even if the projections on each of the remaining nine final-time SVs are small compared to the projection on the leading SV at final time, they do contribute to increasing the total projected part: the SV 10-dimensional subspace accounts for over 40% of the average CRCM IV at final time. It was found that the total perturbation energy is generally dominated by the slow-developing perturbations that are not projected into the first ten SVs. Nevertheless, the IV growth itself is represented in a large proportion by the 10-dimensional subspace of SVs, which accounts for up to 70% of the CRCM IV growth in thirty-six hours.

In this study, we continue the comparative analysis focusing on the CRCM-IV total-energy tendency equation integrated over the free domain and expressed as in Eq. 3.12. In these conditions, the IV total tendency equals the IV sources and sinks within the free domain. The fact that the growth of CRCM IV corresponded in a large proportion to the growth of the most unstable linear perturbations suggests the use of the leading SVs in the representation of the sources of IV.

Figure 3.6 plots the total tendencies for the next six hours corresponding to the CRCM-IV total field and to the projected part on the first SV and on the ten SVs from *SetA*, together with the CRCM IV local variation. If the sources of IV are indeed represented by the most unstable linear perturbations, then the total variation for the total field must be close to the total variation for the projected part in the directions of the set of SVs. We note that the leading SV explains a large part of the local tendency for the next six hours, but it is not sufficient to explain the entire tendency. When the first ten SVs are considered, the tendency

of the projected part approaches closely the CRCM local tendency. There is still some portion in the total tendency that is not completely explained, especially after twenty-four hours of TLM integration. This indicates that more than ten SVs are needed in order to completely explain all contributions to the IV growth or that nonlinear processes may become important at this stage of growth. This result is in accord with our previous study that showed that on average 70% of the IV growth in thirty-six hours was explained by the first ten SVs. Diaconescu *et al.*, (2011) showed that only a very small part of the CRCM perturbations projects on the ten SVs at initial time. This can be also seen in Figure 3.6, which reveals that, at initial time, a very small part of the sources of IV are explained by the ten leading SVs. The rest of the IV sources of growth come from old perturbations that has already grown. Diaconescu *et al.*, (2011) also showed that during the next thirty-six hours, the projected part evolved from small to large scales and their growth rates decreased, approaching that of the non-projected part. Hence, as shown in Figure 3.6, after twenty-four hours the fraction of IV sources explained by the SVs decreases, possibly indicating that nonlinear processes become important for large perturbations. It is also possible that some physics are missing in the TLM used in the SVs computation.

Because SVs are computed using an OTI of thirty-six hours, it is interesting to compare the total tendencies for the next thirty-six hours too. In this case, the local variation (in $J\ m^{-2}h^{-1}$) is computed as the difference between two moments at 36-hour interval. The divergence term is also expressed in $J\ m^{-2}h^{-1}$ and it is computed as the average value over the 36-hour interval. The total tendency is again computed as the sum of the local tendency and the divergence term. Figure 3.7 shows the total and local tendencies for the next thirty-six hours for the CRCM-IV total field and the total tendency for the projected part on the leading SV and on the first ten SVs from *SetA*. As in the case of 6-hour tendency, the projected part on the first ten SVs completely represents the local tendency but do not explain the entire tendency. The ratio between the total tendency of the CRCM total field and of the projected part on the leading SVs is 0.38, while the ratio between the total tendency of the CRCM total field and of the projected part on the first ten SVs is 0.73. We can conclude that 73% of the sources of IV are represented by the first ten SVs. The small difference from the previous study is due to the fact that now the analysis is restricted to a smaller domain (96 x 96 grid

points) and the total energy accounts only for the kinetic and potential terms that are integrated over all model vertical levels.

3.3.2 SetB1 to SetB10 of SVs and the IV tendency

SetA of SVs focuses on the beginning of a period with rapid growth in the CRCM IV. We investigate further whether the SVs can explain the IV growth in periods with smaller growth. Thus, another ten sets of SVs, *SetB1, SetB2, …, SetB10* are used to study the sources of IV for ten 36-hour periods during the first twelve days of December 1992. As for *SetA*, these sets of SVs have an optimization time of thirty-six hours and a total-energy final norm localized over an area in the free CRCM domain. However, for these sets, the total-energy initial-time norm is global and the sets contain only the first five SVs, which are archived only at initial and final time.

Because SVs are computed with a global initial-time norm, not all of them are located completely into the CRCM domain at initial time, and therefore not all SVs must be counted as possible sources of CRCM IV. An example is given in Figure 3.8, which displays the five initial-time SVs from *SetB7* in terms of temperature at the model eta level 0.46. Figure 3.8 highlights that the first two initial-time SVs are located in a large proportion out of the CRCM free domain, which is delimitated in figure by the magenta line. Hence, only SVs no 3, 4 and 5 are considered as valid. However, there are other sets of SVs for which all five SVs are situated in a large proportion into the CRCM free domain, as can be seen in Figure 3.9, which displays the temperature fields for the five initial-time SVs of *SetB8*.

Consequently, we have proceeded to the selection of those initial-time SVs that are located mostly within the CRCM free domain. Table 3.2 presents the amplification factors (Eq. 3.16) for the selected SVs. These factors have been recalculated after the SV interpolation over the CRCM grid using a norm defined over the CRCM domain of analysis. The table allows to identify the fastest "free-domain" SV for each set of SVs. We note that SVs are issued from computation in decreasing order of their amplification factors, which were computed using the global total energy norms at both initial and final time. After the interpolation over the CRCM domain and the recompilation with CRCM-restrained norms, this order is broken for some SVs. For eight of the ten sets of SVs, the first SV located into

the free domain is the fastest SV. For *SetB6* and *SetB8* the fastest growing SV is not the first SV. However, in theses two cases the fastest SV and the first SV have very close amplification factors.

To eliminate the confusion, we reordered the "free-domain" SVs according to the amplification factors computed over the CRCM free domain. The resulted spectra of amplification for all ten sets of SVs are plotted in Figure 3.10. Only four sets of SVs have kept their five SVs, two sets now have four SVs and four sets now have only three SVs. For the most part of the sets, the amplification factors of the leading SV have values between 100 and 150. A particular case is *SetB8* for which the amplification factor of the leading SV reaches the value of 272.

In the following, we have proceeded to the projection of CRCM perturbations on the bases of SVs. Because the total dry energy norm was used in the computation of SVs, the projection on the initial SVs in very small in all ten cases. As Diaconescu *et al.* (2011) noted, this is due to the fact that the total-energy SVs are outside the attractor of the system at initial time. However, the SVs rapidly rotate toward the attractor and the projection on the evolved SVs is generally highter. Figure 3.11 shows the total and local tendencies in the next thirty-six hours for the CRCM-IV total field together with the total tendencies for the projected part on the leading SV and on all SVs from the ten sets of SVs. Because the projection on initial-time SVs is very small, the figure illustrates the projection onto the final SVs. We note that the CRCM tendencies are computed at 6-hour interval, while the projected-part tendencies are computed only at intervals of 24 hours corresponding to the times when the sets of SVs are computed. The sources of IV have a primary maximum between 3 December at 12:00 UTC and 6 December at 12:00 UTC, and the other three secondary maxima centred on 2 December at 06:00 UTC, 8 December at 00:00 UTC and 11 December at 12:00 UTC. The projected part on all SVs, presents also a primary maximum between 3 December at 12:00 UTC and 6 December at 12:00 UTC as does the CRCM total field. However, the secondary maximum observed on 8 December at 12:00 UTC does not have any equivalent in the CRCM total field. Concerning the projection on the leading SV, only one set of SVs presents a significant projection: *SetB4* with ICs from 4 December at 12:00 UTC. From 5 December at 18:00 UTC to 10 December at 12:00 UTC, the local variation of CRCM IV has generally

negative values, which indicates a local decrease in IV. For this period, the dominant role is played by the export of IV out of the free domain and as a consequence, the projection into the SV truncated bases is rather small. From 1 December at 12:00 UTC to 5 December at 18:00 UTC, the local variation has positive values indicating a local growth of the CRCM IV. The projection on the SV truncated bases does not explain completely the sources of growth but they are important. For 3 December at 12:00 UTC, the CRCM projected-part total variation in the next thirty-six hours approaches the CRCM total tendency.

The average fraction of CRCM-IV sources explained by the SVs is given by the ratio between the total tendencies of the projected part and of the total CRCM field. Figure 3.12 shows that for the first 5 days of December, more than 36% of the sources of instability are represented by the four to five fastest SVs computed with the ICs from the CRCM basic state. On 4 December at 12:00 UTC, the total variation due to the sources of instability has a large value, while the export out of the domain is relatively small (see Figure 3.11). 51% of the IV sources are represented by the first SV, while the other three SVs from *SetB4* increase the percentage to 56%. Therefore, in this case, the leading SV has the dominant role. The *SetA* of SVs was computed from the same ICs as the *SetB4* but it uses a localized initial time norm and ten SVs. The ten SVs explained up to 73% of the instability sources. This suggests that a larger number of SVs will explain a larger fraction of the sources of instability.

Figure 3.11 shows that on 3 December at 12:00 UTC, the local tendency is relatively small. The local tendency is half of the total variation because an approximately equal part is exported out of the domain. Figure 3.12 shows that in this case, 85% of the sources of instability are represented by the first five SVs. However, only 14% of the sources are represented by the leading SV, the largest projection being realised on SV no 3 (not shown). On 1 December at 12:00 UTC and 2 December at 12:00 UTC, the leading SV has also a small contribution. These results highlight that the leading SV does not have the dominant role in explaining the source of instability for cases with relatively small tendencies.

For the period of 5 December 12:00 UTC until 10 December 12:00 UTC, the projection is less than 20%. Although there are some sources of instability within the domain, this period is dominated by the IV export out of the domain. As a result, the local tendency of IV decreases for the next thirty-six hours. We must keep in mind that in this study we are far

away from the ICs. As a concequence, the CRCM perturbations, at one particular instance, are:

- perturbations that "will grow" and that can be counted as sources of IV;
- perturbations that are in "a growing state" and that can also be counted as sources of IV;
- perturbations that have "already grown", which are only transported;

Once arrived close to the lateral boundary of the analysis domain, these perturbations will be exported outside of the analysis domain. SVs are developed into the assumption of small linear perturbations with rapid growth and they can explain only the initial growth of atmospheric perturbations. Therefore, only the perturbations that "will grow" can be captured by the most unstable SVs. The fact that the sources of instability in this period do not project into the directions of the fastest SVs suggests that they may be already in "a growing state" and having important nonlinear characteristics, or that they have a smaller growth rate suggesting the need of SVs with smaller growth rate in order to decompose them.

The projection on the leading SVs noted in the case of 4 December at 12:00 UTC brings up the question: Can the leading SV be used to predict periods with growth or decay in the CRCM IV? To answer that question, the CRCM IV local variation for the next thirty-six hours is compared with the tendency of the leading SV from *SetB1, SetB2, ..., SetB10* in Figure 3.13. The maximum CRCM tendency of growth for the next thirty-six hours is reached on 5 December at 00 UTC. A secondary maximum is observed also on 2 December at 06:00 UTC. For the period of 5 December at 18:00 UTC until 10 December at 18:00 UTC, the CRCM IV has negative local tendency. SV tendency also oscillates and presents minima and maxima that do not always correspond with those of CRCM. The main maximum of SV tendency corresponds to 8 December at 12:00 UTC, while the main CRCM maximum is developed between 4 December at 12:00 UTC and 5 December at 12:00 UTC. As mentioned previously, the leading SV from *SetB8* is characterised by an amplification much larger than the other sets of SVs. Hence, for 8 December at 12:00 UTC the first SV indicates the possibility of perturbation linear growth that is not present in the CRCM case. Perhaps nonlinear processes dominate the local growth in this period in the CRCM case or some physics are missing in the TLM used in the SVs computation. These sources of growth are

compensated by the export out of the free domain (see figure 11); as a consequence, the CRCM total perturbation energy will not change in the next thirty-six hours following 8 December at 12:00 UTC. Also, it can be seen that the SV tendency for *SetB4* is only slightly larger than that of *setB2* and it does not anticipate a higher growth in IV than usual.

We therefore conclude that for the set-up used in this analysis, the leading SV alone does not permit to anticipate periods with a higher IV growth than usual.

3.4 Summary and concluding remarks

In this paper, we analyzed the IV in an ensemble of simulations carried out with the Canadian RCM version 5 for the first two weeks of December 1992, in terms of total perturbation energy of an ensemble of twenty-one simulations. The simulations are centred on the North American continent and have the same model configuration except the ICs that correspond to different starting dates between October 23 and November 12 at 00:00 UTC. The basic hypothesis of this study is that the IV is arising through rapidly growing perturbations develop in dynamically unstable regions, which are further transported by the basic state, and will finally be forced to disappear when they approach the model lateral boundaries (MLB).

We focus on the IV total-energy tendency and consider that the IV sources are given by the hydrodynamic instabilities while the main sinks arise from the MLB forcing. To eliminate the sinks of IV caused by the MLB forcing, we restrained the analysis domain to the central portion of the regional domain. A consequence of reducing the domain of analysis is the appearance of a horizontal transport of IV out of the free zone into the buffer zone. Hence the IV local variation will be the result of the competition between the sources of instability (the total variation) and the IV export out from the free domain.

Our study shows that in periods with an important IV export the IV local variation has negative values corresponding to a reduction of the IV in the analysis domain. The periods with positive IV local variations are dominated by the sources of instability, the export having smaller values.

Another issue analysed in this paper involves the proportion in which the sources of growth in the CRCM IV can be explained by the leading singular vector (SV) or by a truncated base with a small number of SVs. SVs capture the dynamically most unstable perturbations and finite-time instabilities into the context of a linear model named the tangent linear model (Buizza, 1994). By definition, SVs form an orthogonal set of perturbations that, according to linear theory, will grow fastest over a finite-time interval with respect to norms specified at initial and final times.

This study is a follow-up of a previous paper (Diaconescu et al., 2011) where one set of SVs was used to explain the growth observed in the CRCM IV for the period of 4 December 12:00 UTC until 6 December 00:00 UTC. That previous study, which focused on a period with rapid growth in the IV, has shown that an average of 70% of the CRCM IV growth in thirty-six hours was explained by the first ten SVs computed with the global GEM model using ICs from the CRCM basic state.

The present study goes beyond by investigating whether the SVs can explain the variations in the CRCM IV in periods with smaller growth. Another ten sets of five SVs were used to measure the sources of IV for ten 36-hour periods extended on the first twelve days of December 1992. They use as ICs the CRCM reference state embedded into the ERA40 fields and they are computed to maximise the total energy over an area located into the CRCM free domain after thirty-six hours of global GEM linear integration. The tendencies of the CRCM projected-part on the leading SV and on every SVs from the ten sets of SVs are compared with the total tendency of the CRCM total field. If the sources of IV are indeed represented by the most unstable linear perturbations, then the total variation for the total field should be identical to the total variation for the projected part in the directions of the set of SVs.

For the first 5 days of December, more than 36% of the IV growth is represented by the four (or five) fastest SVs computed with the ICs from the CRCM basic state. This period was dominated by the sources of instability, the IV export out from the analysis domain being relatively small. Only in the case of 4 December at 12:00 UTC was the projection dominated by the leading SV: 51% of the IV sources are represented by the first SV, while the other three SVs increase the percentage to 56%. When a set of ten SVs is used, the percentage of projection grows to 73%. The paper also shows that for periods with a smaller IV tendency,

the leading SV does not have the dominant role in explaining the sources of instability.

Although there are some sources of instability in the domain, relatively small projections are obtained in periods where the IV local tendency is dominated by the IV export out of the analysis domain. The fact that the source of instability in this period does not project into the directions of the fastest SVs suggests that they may be already in "a growing state" and having important nonlinear characteristics, or that they have a smaller growth rate suggesting the need of SVs with smaller growth rate in order to decompose them.

This work has shown that for the set-up used in this analysis, the leading SV alone does not permit anticipating periods with higher IV growths than usual.

One technical limitation of this study was the use a global model to compute the SVs. To focus on the region of interest, we restricted the domain of the final-time norm. Ideally, we would have used SVs issued from a limited-area TLM, which would take into account effects of the driving boundary conditions on the growing disturbances of the CRCM. However, this study brought some valuable information. It has been shown that a varying number of SVs is required to explain the sources of IV, and categorically one SV is insufficient for explaining the IV tendency. This is due to the fact that the small sources of instability expressed by SVs are arising from a mass of mature perturbations that are already in "a growing state" and have smaller growth rate or even important nonlinear characteristics. Sometimes, the linear growth of the small instabilities is so high that it will dominate the growth associated with old perturbations located in the domain. However, beyond the initial period of linear growth, the nonlinear processes will dominate the IV tendency. As a consequence, there are periods in the IV evolution when a small number of SVs suffices to explain a large part of the IV tendency; but there are also times when the nonlinear processes are dominant. It is important to note that, at these particular times, the export of IV out of the analysis domain exhibited large values; this suggests that the CRCM perturbations approached the MLBs and therefore they had the time to grow. SV growth is based only on the interaction of the perturbation with the basic state, which feeds it and does not consider any interaction between perturbations at different states of their life cycle. This kind of interaction is the source of nonlinearities. It is therefore of great interest to study the role of such interactions in the IV evolution.

Also, another possibility is to use bred vectors (Kalnay, 2002) instead of SVs. The breeding method consists of generating a control run of a nonlinear model for a short period of time (e.g., 6 h), then perturbing the atmospheric initial conditions and running the same model again for the same period of time (perturbed run). The difference between the two model runs at the final time is adjusted to the amplitude of the initial perturbation and is added to the new control initial condition. The method is then repeated. After a few days of iteration, the difference between the control and perturbed model runs represents a sample of fast-growing nonlinear perturbations (e.g., Szunyogh *et al.*, 1997; Kalnay, 2000). Hence, bred vectors represents finite-amplitude, finite-time vectors that do not require linearizing the model, are independent of the norm used to define the size of the perturbation and can be computed exactly for the regional domain.

Acknowledgements

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Graduate Scholarship (CGSD3), as well as by the CRCMD Network supported by the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS) and the Ouranos Consortium. Ouranos and the ESCER Centre provided the computing resources. We thank Prof. Eugenia Kalnay and Prof. Pierre Gauthier for their comments on a previous version that helped improve the manuscript.

Table 3.1: Configuration of the sets of SVs

Name of SV set	Date of ICs	No of SVs	Initial Norm
SetA	4 December 1992 at 12 UTC	10	localized
SetB1	1 December 1992 at 12 UTC	5	global
SetB2	2 December 1992 at 12 UTC	5	global
SetB3	3 December 1992 at 12 UTC	5	global
SetB4	4 December 1992 at 12 UTC	5	global
SetB5	5 December 1992 at 12 UTC	5	global
SetB6	6 December 1992 at 12 UTC	5	global
SetB7	7 December 1992 at 12 UTC	5	global
SetB8	8 December 1992 at 12 UTC	5	global
SetB9	9 December 1992 at 12 UTC	5	global
SetB10	10 December 1992 at 12 UTC	5	global

Table 3.2: Growth rate for the SVs that are situated mostly in the free zone at initial time with the most rapid SV in red; growth is computed with norms over the CRCM free domain

	Date of ICs	SV1	SV2	SV3	SV4	SV5
SetB1	1 December 1992 at 12 UTC	**132.20**	125.14	82.45	69.15	73.09
SetB2	2 December 1992 at 12 UTC	**134.14**	133.73	97.24	73.16	76.11
SetB3	3 December 1992 at 12 UTC	**114.86**	111.23	60.69	63.87	78.17
SetB4	4 December 1992 at 12 UTC	**143.28**	--------	94.92	93.41	90.81
SetB5	5 December 1992 at 12 UTC	**145.55**	124.39	95.13	71.42	--------
SetB6	6 December 1992 at 12 UTC	91.59	89.83	---------	**92.60**	--------
SetB7	7 December 1992 at 12 UTC	-------	---------	**111,42**	111.12	88.51
SetB8	8 December 1992 at 12 UTC	268.62	**272.49**	134.96	93.90	87.17
SetB9	9 December 1992 at 12 UTC	-------	---------	**143.94**	105.58	133.21
SetB10	10 December 1992 at 12 UTC	**165.42**	158.11	128.22	----------	--------

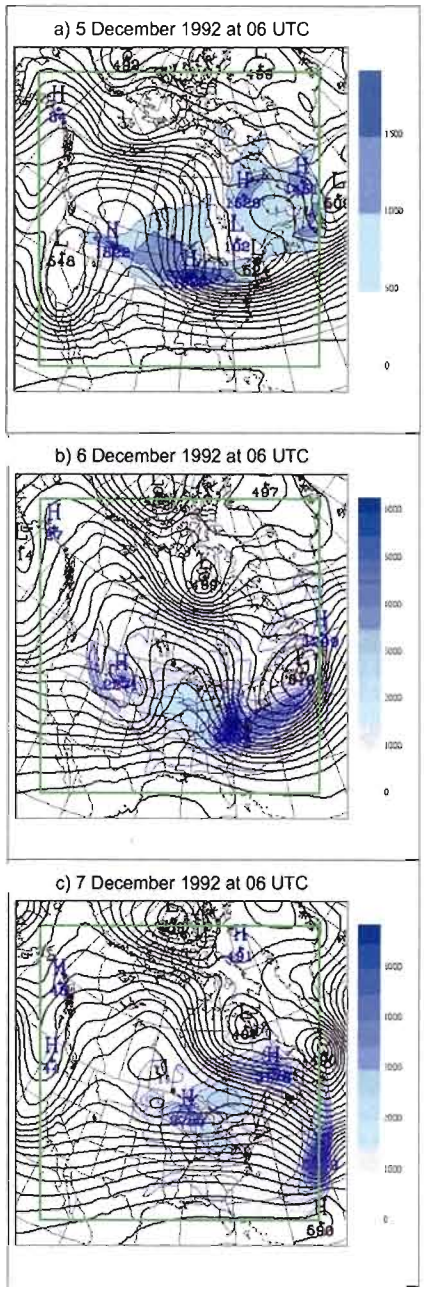


Figure 3.1: 500 hPa geopotential field (*in dam*) for the reference simulation and the average CRCM-perturbation total energy (*in* $kJ\,m^{-2}$) integrated between 100 hPa and surface. The geopotential field (*black contour*) is contoured at 4 dam intervals, while the CRCM-perturbation total energy is plotted in colours. The geen square excludes the buffer zone of ten points.

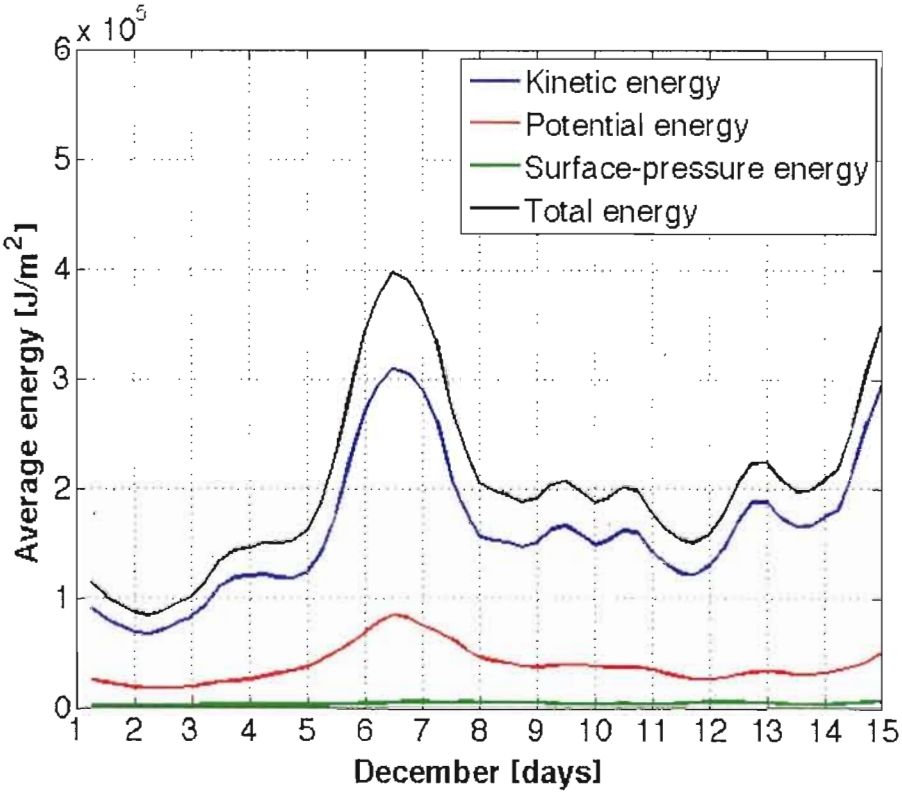


Figure 3.2: Time evolution over the first two weeks of December 1992 for CRCM (*black line*) total average energy and its components: (*blue line*) kinetic, (*red line*) potential and (*green line*) surface-pressure energy. All terms are integrated over the entire CRCM domain.

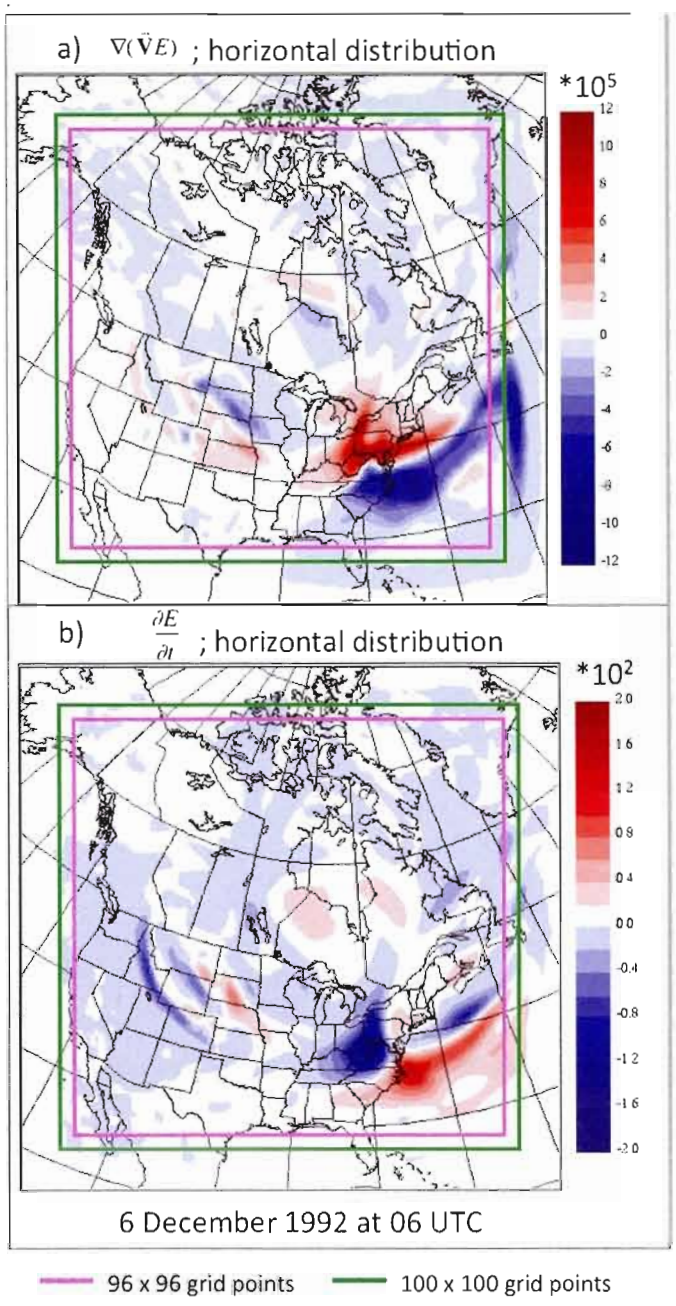


Figure 3.3: Horizontal distribution of (a) the perturbations total-energy divergence term ($\nabla \cdot (\mathbf{V}E)_h$) and (b) the perturbations total-energy local tendency for the next 6 hours ($\frac{\partial E_h}{\partial t}$) on 6 December 1992 at 06UTC. The geen square excludes the buffer zone of ten points, while the magenta square shows the analysis domain.

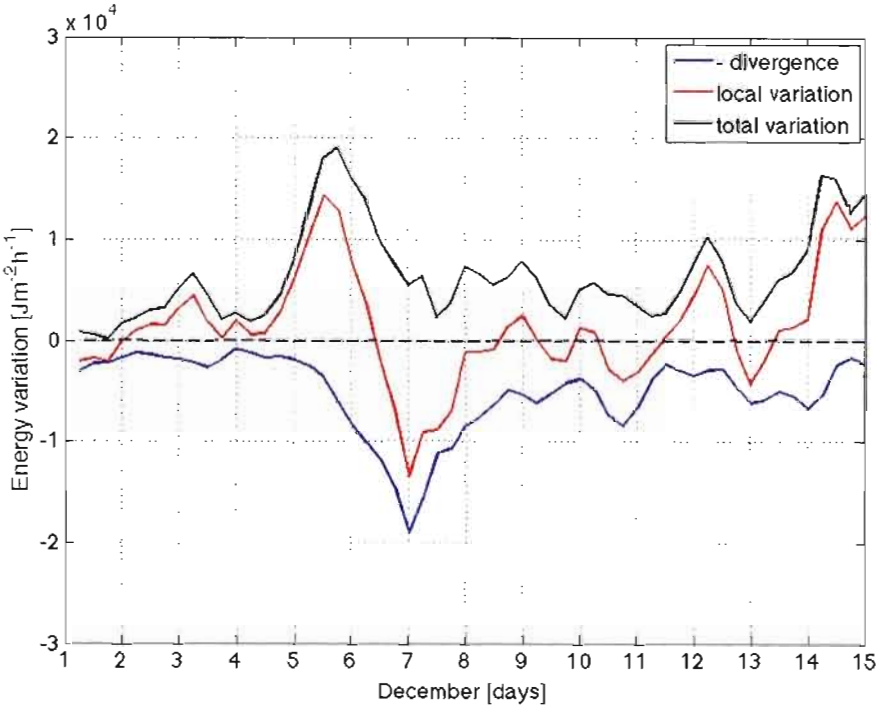


Figure 3.4: Equation for CRCM IV total tendency for the next 6 hours. The local tendency is represented in red, the IV export out of the analysis domain or the 6-hour average IV divergence term with negative sign is showed in blue, while the IV total variation for the next 6 hours or the sources is plotted in black.

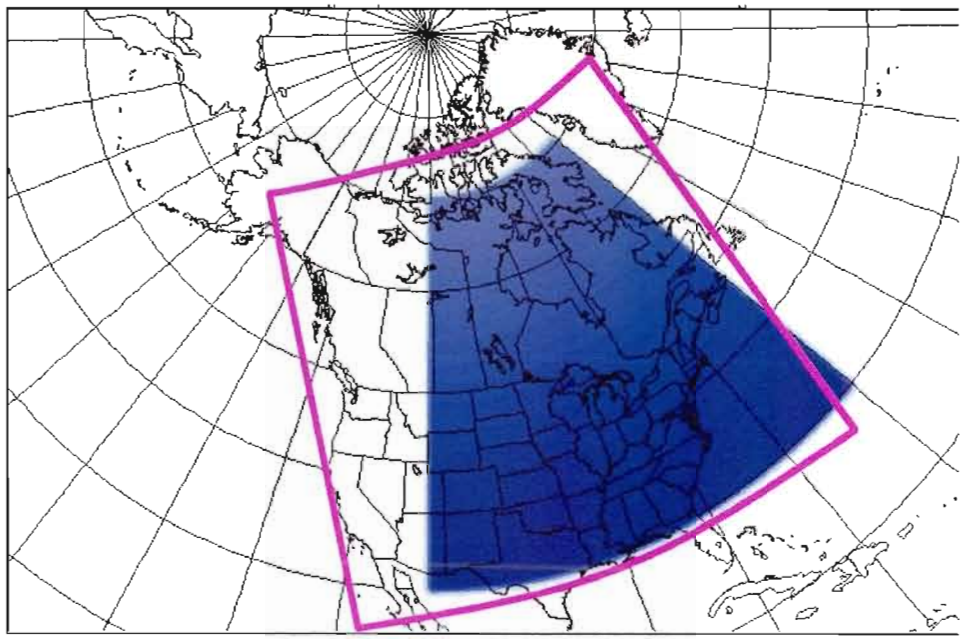


Figure 3.5: The initial- and final-norm domain for *SetA* of SVs and the final-norm domain for the sets *SetB1*, *SetB2*, ..., *SetB10*. The magenta line delimitates the CRCM analyses domain.

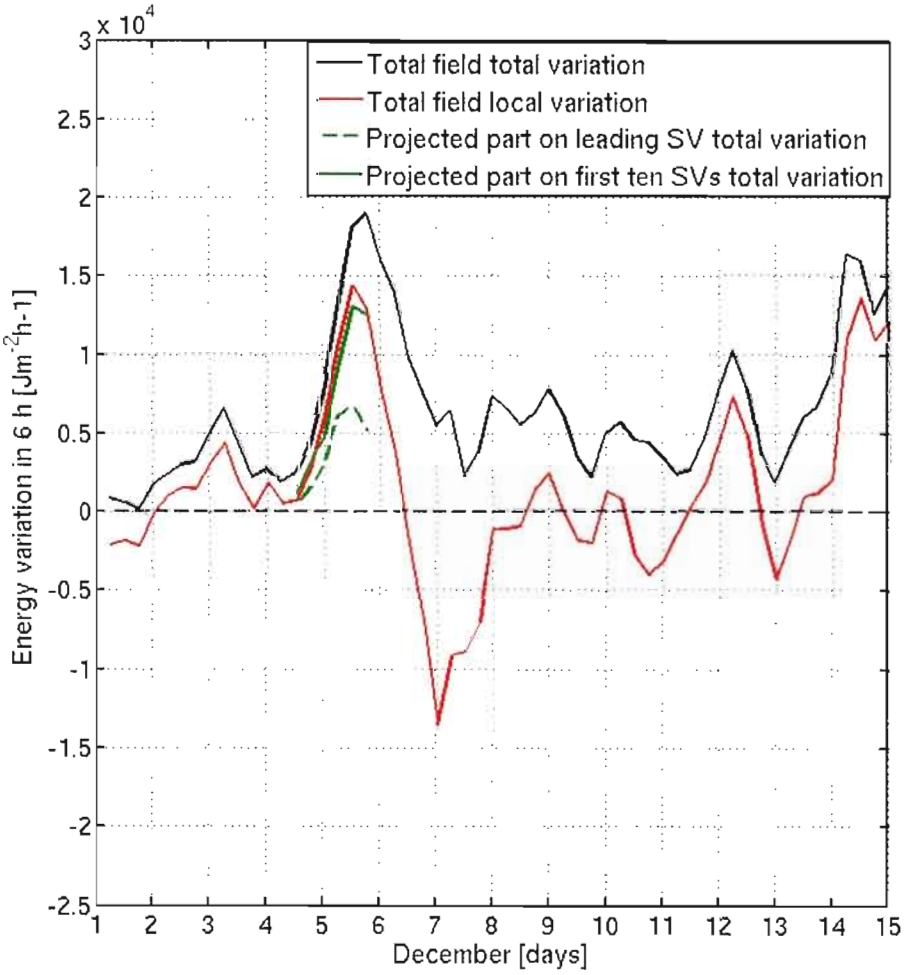


Figure 3.6: Total tendencies for the next 6 hours corresponding to (black line) the total CRCM IV and to the (green line) projected part on the (dashed line) first SV and the (solid line) ten SVs from *SetA*. The red line displays the CRCM IV local variation for the total field.

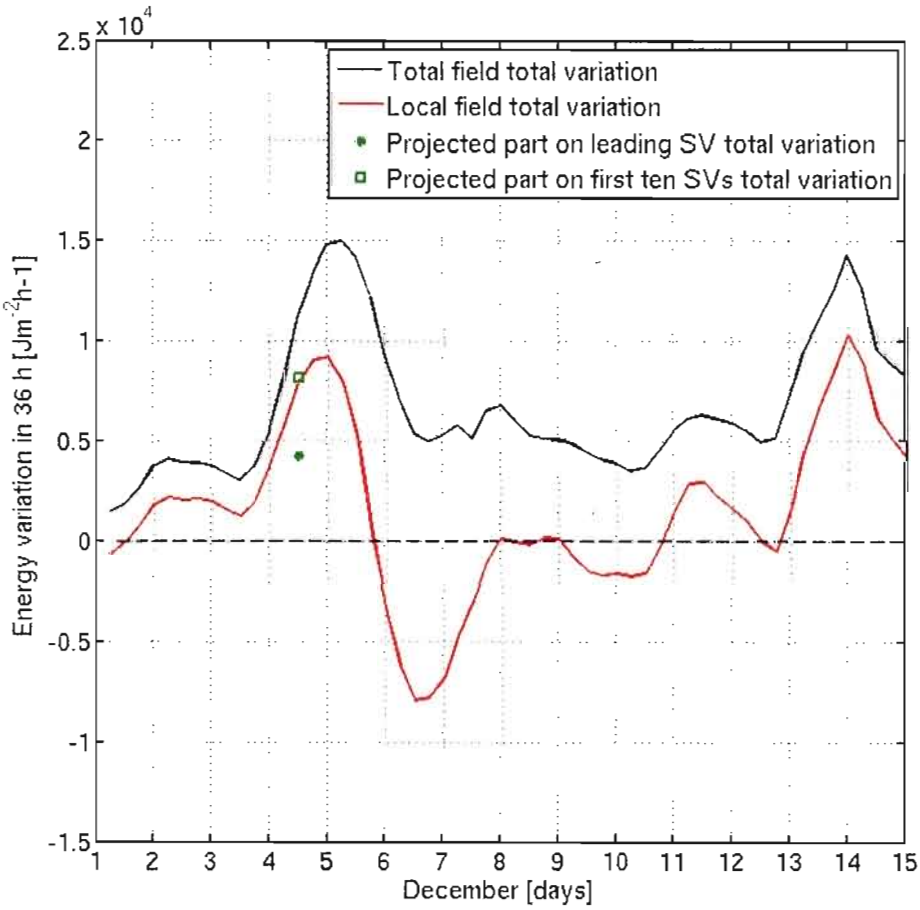


Figure 3.7: Total tendencies for the next 36 hours corresponding to (black line) the total CRCM IV and to the projected part on the (green star) first SV and the (green square) ten SVs from *SetA*. The red line displays the CRCM IV local variation for the total field.

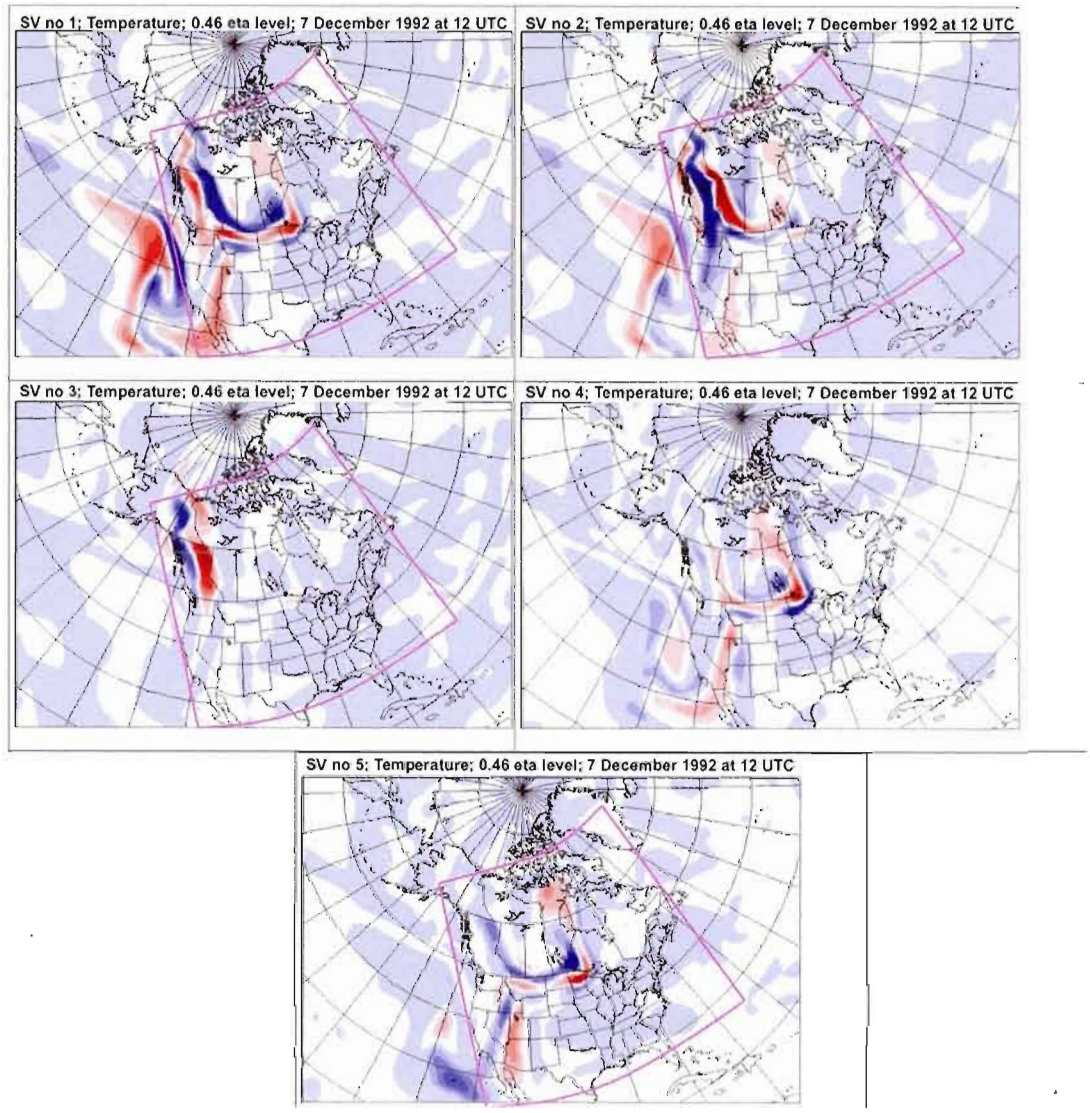


Figure 3.8: Temperature fields for the five initial-time SVs from *SetB7* at the model eta level 0.46.

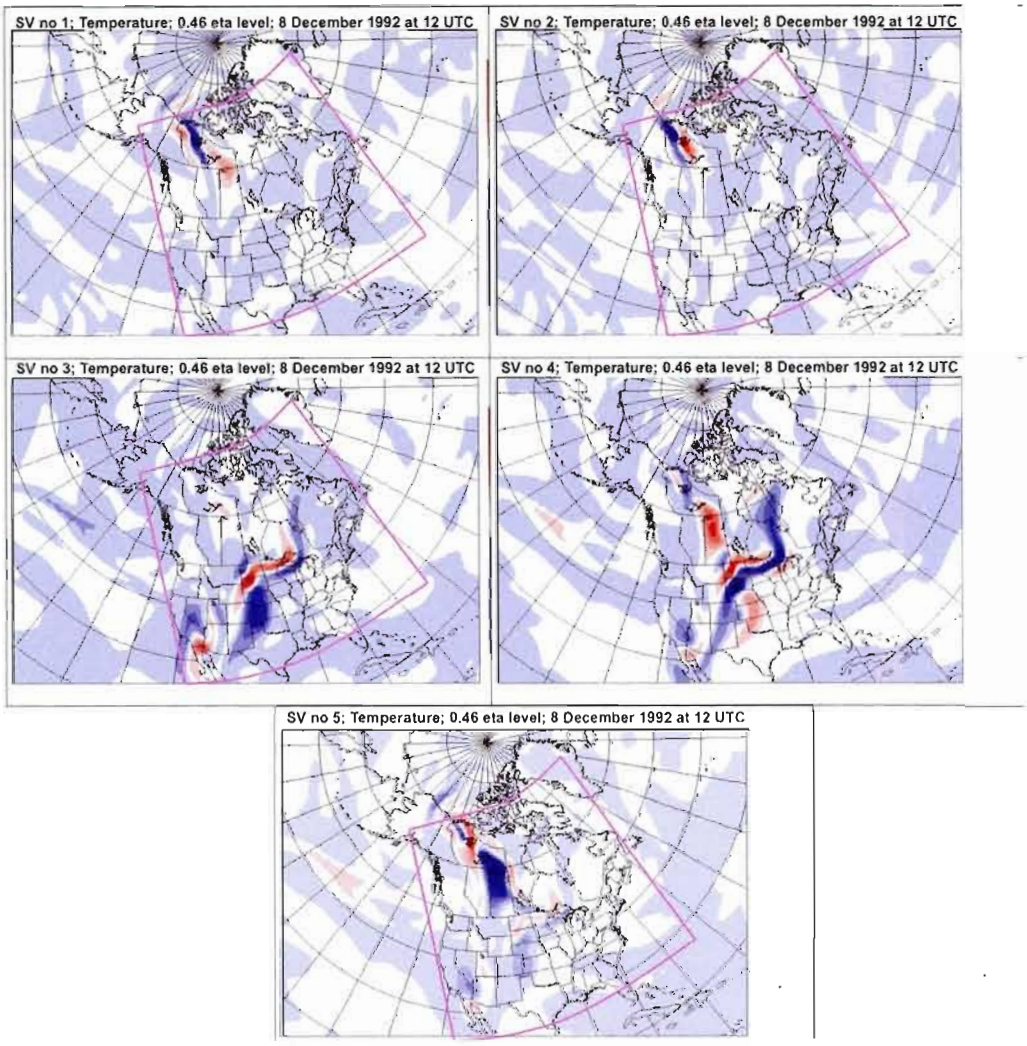


Figure 3.9: Temperature fields for the five initial-time SVs from *SetB8* at the model eta level 0.46.

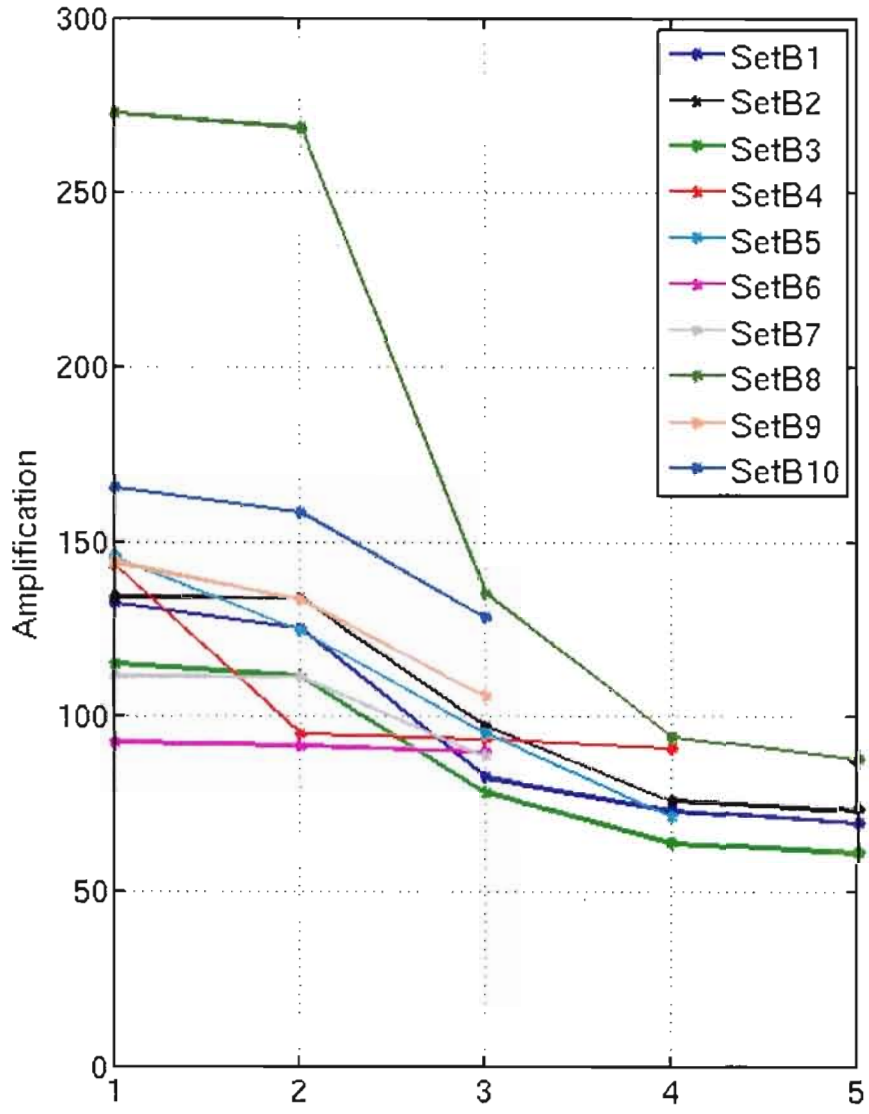


Figure 3.10: Amplification factors for the ten sets of SVs.

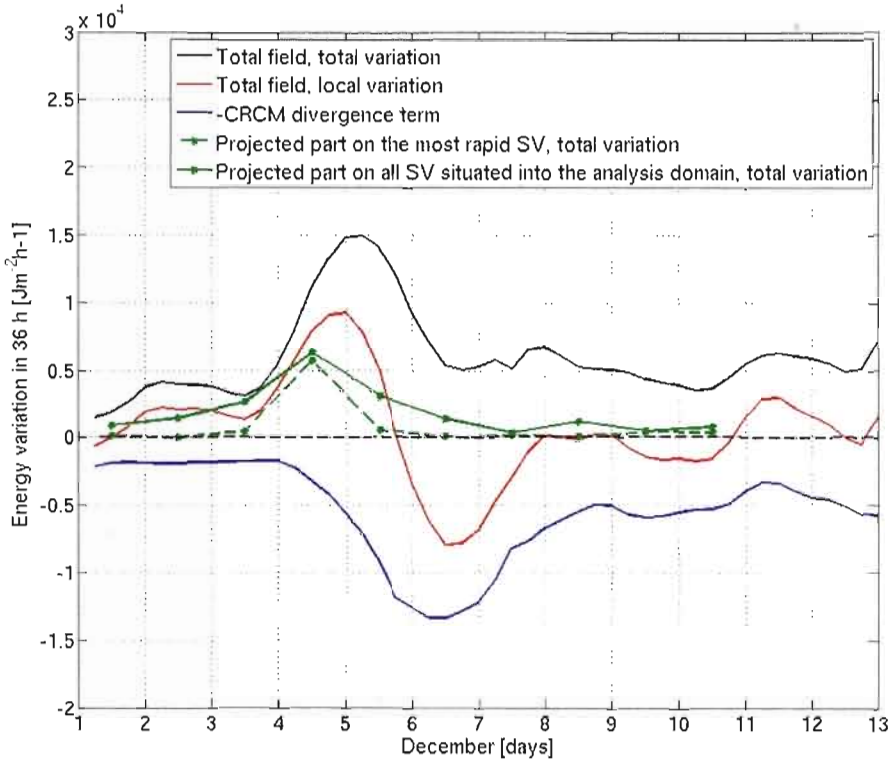


Figure 3.11: Total tendencies for the next 36 hours corresponding to (black line) the total CRCM IV and to the projected part on the (green dashed line) first SV and the (green solid line) ten SVs from *SetA*. The red line displays the CRCM IV local variation for the total field and the IV export out of the analysis domain or the 36-hour average IV divergence term with negative sign is showed in blue.

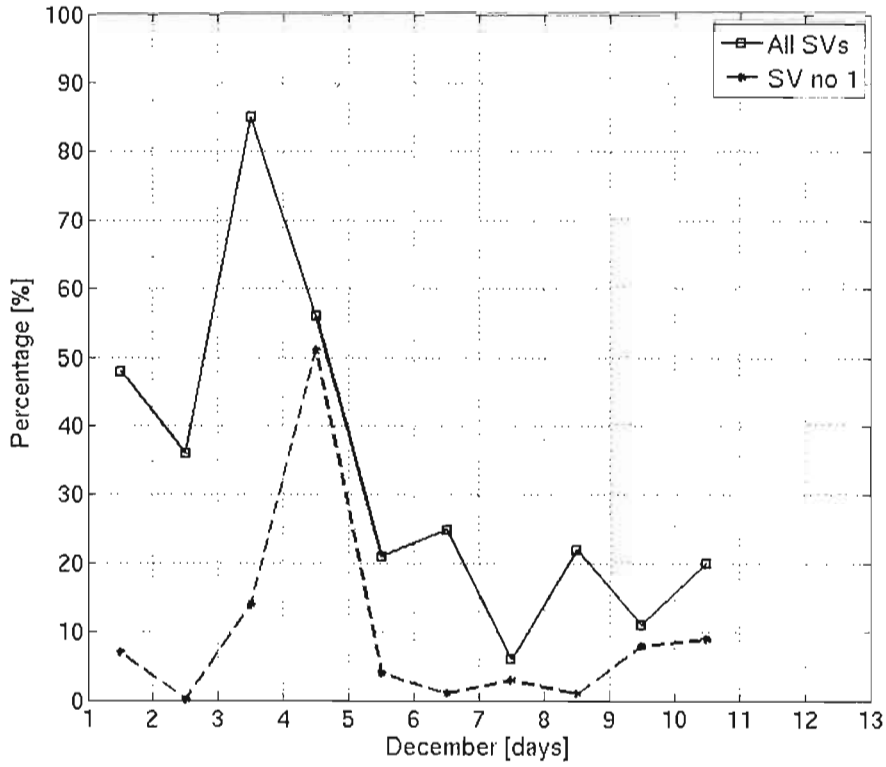


Figure 3.12: The average fraction of CRCM-IV sources explained by the (dashed line) leading SV and (solid line) all SVs of the sets noted *SetB1, SetB2, ..., SetB10.*

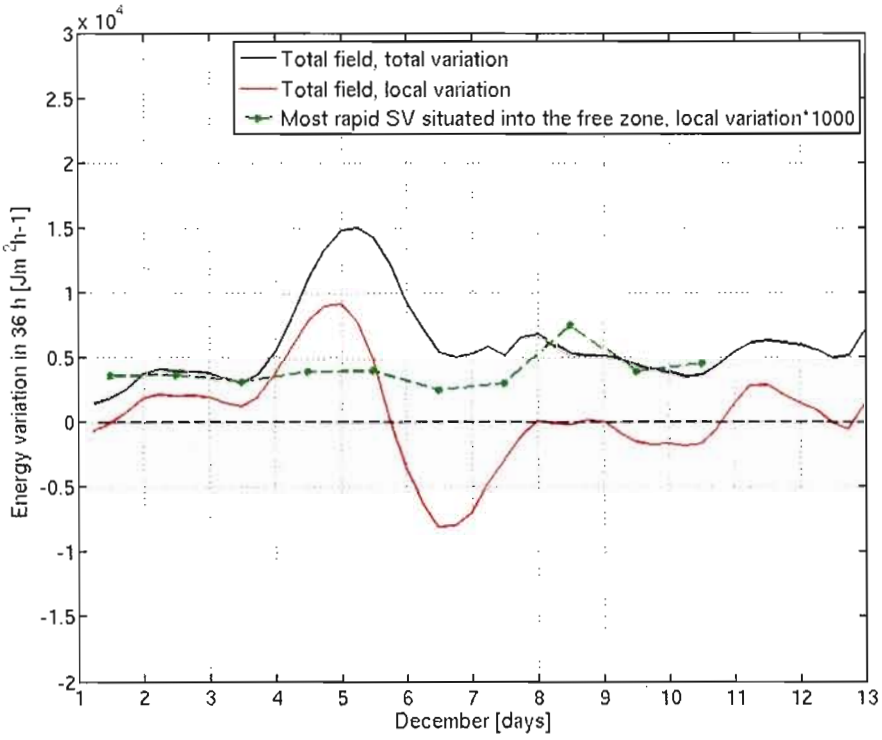


Figure 3.13: Local tendencies for the next 36 hours corresponding to the (red line) CRCM IV and to the (green stars) leading SV from *SetB1, SetB2, ..., SetB10*. The black line displays the CRCM IV total variation for the next 36 hours. For clarity, the SV tendency has been multiplied by 1000.

CHAPITRE IV

CONCLUSION

4.1 Conclusion principale de la recherche

La variabilité interne (VI) représente une question importante en modélisation régionale du climat. La VI est définie comme la dispersion des solutions d'un ensemble de simulations causée par l'amplification de petites différences présentes dans les conditions initiales (CI). Plusieurs études ont montré que l'évolution temporelle de la VI d'un modèle régional du climat (MRC) a des périodes caractérisées par de fortes augmentations et des périodes avec des valeurs plus faibles. Les maxima de VI sont en général associés à des périodes et régions avec d'intenses activités synoptiques.

Ce projet avait pour but principal de vérifier l'hypothèse selon laquelle les maxima de VI sont dus à la croissance rapide des perturbations développées dans des régions atmosphériques instables. L'élément déclencheur est représenté par les instabilités hydrodynamiques telle l'instabilité barocline ou barotrope. Pour décrire l'espace instable des perturbations, nous avons fait appel à la technique des vecteurs singuliers (VS). Cette technique est utilisée en particulier dans la prévision numérique du temps pour générer les CI dans les systèmes de prévision d'ensemble (Molteni *et al.*, 1996; Palmer *et al.*, 2007; Yamaguchi and Komori, 2009). Cependant, c'est la première fois qu'elle est utilisée dans l'analyse de la VI d'un modèle du climat. Les VS permettent d'identifier la série des perturbations avec la plus grande croissance dans la version linéarisée du modèle et dans un intervalle de temps limité. La croissance est mesurée à l'aide des normes définies au début et à la fin de l'intervalle de temps appelé temps d'optimisation (OTI).

La première partie de la thèse avait comme objectif de déterminer quelles étaient les principales caractéristiques des VS. La recherche bibliographique présentée dans le deuxième

chapitre de la thèse a montré que, en général, la structure et les propriétés des VS dépendent de la norme au temps initial, la complexité et la résolution du modèle linéaire tangent (TLM) ainsi que des OTI utilisés dans leur calcul. Le plus populaire choix de normes au temps initial et final est la norme de l'énergie totale. La plupart des études utilisent des OTI qui varient en général entre 24 heures et 48 heures. Le choix de l'OTI est conditionné par la validité de l'approximation du TLM qui est fonction des échelles du champ : les petites échelles sont dominées par des processus non linéaires après 12 heures, alors que les grandes échelles sont linéaires pour plus de 48 heures. En général, un OTI de 48 heures et un TLM sec (sans processus humides) avec une faible résolution sont utilisés dans les études des perturbations extratropicales baroclines. Par contre, si les VS visent une région tropicale, un TLM humide avec une haute résolution et un OTI de 24 heures sont préférables parce qu'ils permettent de mieux capter les petites échelles caractéristiques aux processus humides. Le calcul de VS présente un coût numérique important, ce qui fait que dans la plupart des études, les bases des vecteurs sont tronquées pour se limiter aux premiers 10 à 50 VS, donc, aux perturbations les plus instables de l'écoulement.

Ce projet a été consacré à l'analyse de la VI dans des simulations du modèle régional canadien du climat (MRCC) sur un domaine centré sur le continent nord-américain pour une période hivernale (décembre 1992). Pour quantifier la VI nous avons réalisé un ensemble de 21 simulations avec la même configuration et un décalage de 24 heures entre les CIs (de 23 octobre 1992 au 12 novembre 1992). L'ensemble a été décomposé dans un état de référence représenté par l'une des simulations et une série de perturbations par rapport à cette référence. La VI a été décrite en termes d'énergie totale des perturbations. L'énergie totale a été utilisée aussi comme norme au temps initial et final dans le calcul des VS. La norme de l'énergie totale est souvent utilisée pour calculer les VS dans les régions extratropicales parce qu'elle permet de capter la dynamique des instabilités baroclines et barotropes caractéristiques à ces régions.

Dans un premier volet, nous nous sommes intéressés à un épisode marqué par une forte augmentation de la VI (de 4 décembre à 12:00 UTC au 6 décembre à 00:00 UTC) et sa décomposition en VS. Une série de dix VS a été calculée à partir de l'état de référence du MRCC en utilisant la norme de l'énergie totale localisée sur le domaine MRCC. Étant donné

que l'analyse est consacrée à une période hivernale, nous avons décidé d'utiliser un TLM sec et un OTI de 36 heures. L'analyse du spectre des VS a montré que les dix VS avaient de fortes croissances pendant 36 heures, la série étant dominée par le premier VS.

La projection des perturbations MRCC sur la base de dix VS a permis de quantifier dans quelle mesure la VI est expliquée par les dix premiers VS. Nos résultats ont montré que seulement une très petite partie des perturbations MRCC étaient projetée dans les directions des VS au temps initial. Cependant, cette projection augmente substantiellement pendant les 36 heures. L'évolution temporelle des énergies correspondant à la partie projetée et à la partie non projetée a permis de constater que, dans l'intervalle de 36 heures, la partie projetée se développait rapidement pour atteindre une structure de grandes échelles et son taux de croissance diminuait pour approcher la valeur de la partie non projetée. Au temps final, en moyenne, 40% de l'énergie totale des perturbations MRCC était expliquée par les dix VS. D'autre part, le faible taux de croissance de la partie non projetée a indiqué que ces perturbations de grandes échelles sont transportées par l'écoulement de référence, tout en contribuant dans une petite proportion à la croissance dans le temps de la VI. Après 36 heures, en moyenne 70% de la croissance de la VI est expliquée par les dix premiers VS. Nous avons conclu que la croissance de la VI pour cette période était due principalement aux perturbations représentées par les premiers VS.

L'analyse de la projection dans les directions de chacun des VS a montré que la projection était faite avec prépondérance sur le premier VS. La comparaison entre les perturbations MRCC et les VS a illustré la grande ressemblance entre la structure horizontale et verticale des perturbations MRCC et celle du premier VS après 24 à 36 heures d'intégration linéaire tangente.

Dans un deuxième volet, nous nous sommes intéressés à l'analyse de la VI pour des périodes des croissances relativement faibles par rapport au cas précédant. L'analyse a visé la variation totale de la VI. La présence des frontières latérales du MRCC à un fort impact sur l'évolution temporelle de la VI : aux frontières latérales, les champs de température et du vent horizontal simulés par le MRCC ont les mêmes valeurs que les champs du modèle pilote. Par conséquent, la VI converge à zéro dans la zone des frontières. Pour réduire

l'influence des frontières latérales du modèle, le domaine d'analyse a été limité à la partie centrale du domaine.

Nous avons montré que, dans ce contexte, les fluctuations dans la VI sont modulées par les processus générateurs de VI (tel que les instabilités hydrodynamiques à l'intérieur du domaine) ainsi que par le transport de la VI à l'extérieur du domaine d'analyse. Nos travaux ont montré que dans les périodes avec un important transport de VI à l'extérieur du domaine, la VI diminue à l'intérieur du domaine. Cependant, les périodes avec des augmentations de la VI sont dominés par les sources des instabilités. Pour quantifier dans quelle mesure ces sources sont représentées par les perturbations avec la plus forte croissance, nous avons projeté les perturbations MRCC dans les directions des cinq premiers VS et nous avons calculé la variation totale de la VI due à la partie projetée.

Nous avons trouvé que, dans les périodes dominées par les sources d'instabilité, plus de 36% de la croissance de la VI était expliquée par les 4 ou 5 premiers VS. Par contre, seulement dans un cas, la projection a été dominée par le premier VS : c'est le cas du maximum principal observé dans la variation temporelle de la VI. Dans les cas où la VI a des tendances plus petites, ce n'est pas le premier VS qui a le rôle dominant dans la projection.

Nous avons montré également que dans les périodes dominées par le transport de VI à l'extérieur du domaine, la projection sur les VS était très petite malgré la présence des sources de VI. Le fait que ces sources de VI ne soient pas représentées par les VS les plus instables indique qu'un nombre plus grand de VS est nécessaire pour leur décomposition ou qu'elles peuvent être dominées par des processus non linéaires.

Le dernier volet de cette thèse concernait la possibilité d'utiliser le premier VS pour estimer les variations temporelles de la VI d'un MRC. Nos résultats montrent que l'utilisation unique du premier VS ne permet pas d'anticiper des périodes caractérisées par une grande croissance dans la VI du modèle.

4.2 Contribution à l'avancement des connaissances et originalité

La VI représente une question importante dans la modélisation régionale du climat. Plusieurs études ont analysé la VI dans des simulations de MRCs et ont remarqué la présence

d'épisodes de forte croissance. Cependant, très peu d'entre elles sont consacrées à l'analyse des processus physiques responsables de cette croissance épisodique. Pour aborder cette question, nous avons fait appel à la technique de VS et quantifié la VI en termes d'énergie des perturbations MRCC par rapport à une simulation de référence. La technique des VS est connue dans les sciences atmosphériques pour ces applications dans la prévision numérique du temps (ex. elle est utilisée pour générer les CI dans les systèmes de prévision d'ensemble). Cependant, c'est la première fois qu'elle est utilisée dans l'analyse de la VI d'un modèle du climat. Donc, l'originalité principale de cette thèse consiste dans l'application d'une méthode spécifique à la prévision numérique du temps pour expliquer la VI spécifique à un modèle du climat. Dans cette étude, la technique des VS nous a permis d'identifier les plus instables perturbations parmi les perturbations MRCC et de vérifier l'hypothèse selon laquelle la croissance épisodique observée dans la VI est due à des perturbations s'étant développées dans des régions atmosphériques instables et ayant un très grand taux de croissance.

Nous avons montré que cette hypothèse peut expliquer les épisodes avec une très rapide croissance de VI mais elle ne peut pas expliquer toutes les variations notées dans la VI. Nos résultats montrent que le nombre de VS nécessaire pour l'explication des sources de VI varie d'un cas à l'autre. La thèse montre également que le premier VS à lui seul ne permet pas d'anticiper les périodes ayant une forte VI. Nous avons également montré que les sources d'instabilité représentées par les VS se développent dans un environnement de perturbations matures avec d'importantes caractéristiques non linéaires. Parfois, la croissance linéaire de ces instabilités de petite échelle est assez grande pour dominer les autres sources de croissance qui sont aussi présentes dans le domaine. Cependant, une fois la période de croissance linéaire dépassée, les processus non linéaires dominent la tendance de la VI. Par conséquent, il y a des périodes où un petit nombre de VS peut expliquer une grande partie de la tendance de la VI, et d'autres périodes où un très grand nombre de VS est requis pour que le même pourcentage de la tendance soit expliqué. La méthodologie que nous avons développée dans le quatrième chapitre a permis de montrer que ces dernières périodes sont aussi caractérisées par un grand transport de VI vers l'extérieur du domaine d'analyse. Ceci suggère qu'une grande partie des perturbations a eu le temps de croître et de dépasser la période où l'approximation linéaire était valide.

En conclusion, cette thèse a permis de comprendre la manière dont la variation temporelle de la VI est influencée par les instabilités hydrodynamiques ainsi que par le transport de VI à l'extérieur du domaine d'analyse.

Je tiens à préciser que les VS utilisés dans cette thèse ont été calculés au centre de recherche en prévision numérique (RPN) par des chercheurs d'Environnement Canada (Professeur Ayrton Zadra). Ma contribution personnelle consiste à la réalisation de l'ensemble des simulations avec la version 5 du MRCC et à la réalisation des toutes les analyses présentées dans cette thèse.

4.3 Limites de la recherche, travaux futurs et recommandations

À cause de limitations informatiques, la plupart des séries de VS utilisées dans la dernière partie de la thèse étaient limitées aux cinq premiers VS ayant une norme initiale globale. Parce que certains de ces VS étaient placés partiellement à l'extérieur du domaine MRCC, nous n'avons pas tenu compte de ces VS dans les analyses. Par conséquent, les séries contenaient un nombre restreint de VS. L'utilisation des séries avec un plus grand nombre de VS permettrait d'augmenter les pourcentages de projections et pourrait apporter plus de finesse à cette étude. Cependant, nous ne croyons pas que les conclusions seraient différentes.

Une autre limitation de cette étude consiste en l'absence du code pour la computation des VS sur un domaine à aire limitée. Par conséquent, nous avons utilisé dans le calcul des VS un TLM développé pour le modèle mondial GEM et le code de calcul adjacent. Pour trouver les VS qui ont un impact sur le domaine MRCC nous avons utilisé une norme finale localisée à l'intérieur du MRCC et les VS ont été interpolés sur la grille MRCC. Aussi, la différence entre les grilles du TLM mondial et du MRCC fait en sorte que le domaine d'analyse MRCC et le domaine de la norme finale ne sont pas identiques. Parce qu'ils proviennent d'un modèle mondial, l'évolution des VS n'est pas influencée par les frontières latérales du MRCC. L'utilisation des VS calculés pour MRCC tiendrait compte des influences des frontières latérales sur les VS et permettrait une analyse sur tout le domaine MRCC.

Il faut aussi mentionner que la technique des VS a permis de sélectionner parmi les perturbations du MRCC, les perturbations avec la plus grande amplification dans un intervalle de temps fini dans le contexte d'un modèle linéaire. La croissance des VS est basée seulement sur l'interaction entre le VS et l'état de base qui varie dans le temps. Par contre, elle ne tient pas compte des interactions entre les perturbations à différents stades de développement. Il est par conséquent intéressant d'étudier plus en détail l'influence de ce type d'interactions sur la VI.

Il serait également intéressant d'étudier plus en détail les processus physiques responsables de la croissance de la VI via la décomposition de la tendance de l'énergie cinétique et de l'énergie potentielle en plusieurs termes basés sur la covariance des perturbations et les gradients de l'état de base. Pour ce but, nous envisageons une collaboration avec Oumarou Nikiema qui a déjà développé deux équations de tendance pour la température potentielle et le tourbillon relatif dans une étude de la VI sur un cas d'été (Nikiema et Laprise, 2010). Cette future étude vise à expliquer les processus physiques responsables de la croissance de la VI dans le MRCC sans l'utilisation du modèle linéaire tangent et sans les approximations reliées aux interpolations des vecteurs sur la grille MRCC.

ANNEXE

ÉTAPES PRÉLIMINAIRES DANS LE DEVELOPEMENT DE LA METHODOLOGIE ET AUTRES ANALYSES

Tel que mentionné souvent dans la thèse, le calcul des VS n'est pas disponible pour le Modèle Régional Canadien du Climat (MRCC5). Les VS sont calculés à l'aide du TLM et ADM du modèle global GEM par des chercheurs d'Environnement Canada au centre de recherche en prévision numérique (RPN) à Dorval, et nous ont été fournis par Dr Ayrton Zadra. Le défie principal de cette recherche a consisté dans l'adaptation des VS issus de cet outil conçu pour la prévision numérique sur un domaine global, pour des analyses sur le domaine à l'aire limité du MRCC5. Dans cette annexe, je présente un historique des étapes préliminaires qui m'ont permis d'aboutir à la méthodologie employée dans la thèse et un résumé des résultats obtenus lors des analyses antérieures aux celles présentées dans les chapitres de la thèse.

Les VS permettent d'identifier la série des perturbations avec la plus forte croissance, à partir d'un état de référence qui évolue dans le temps, dans la version linéarisée du modèle et dans un intervalle de temps limité. Leur calcul dépende du choix des plusieurs paramètres :

- La version des modèles TLM et ADM;
- Le temps d'optimisation;
- Les conditions initiales;
- L'état de référence qui évolue dans le temps;
- Les normes choisies au début et à la fin du temps d'optimisation.

Dans cette étude, le TLM et l'ADM sont ceux du modèle global GEM. Ils incluent un module de la dynamique linéarisée et un module de physique simplifiée qui peut avoir divers degrés de complexité :

- Incluant seulement la diffusion verticale;
- Incluant un module de diffusion horizontale;

- Incluant la précipitation stratiforme;
- Incluant la précipitation convective.

Zadra et al. (2004) présentent une analyse détaillée de l'impact des différents modules physiques simplifiés sur les VS calculés avec le modèle global GEM dans des régions extra-tropicales. Le temps d'optimisation pour les VS calculés dans des régions extra-tropicales peut varier entre 24 heures et 48 heures.

J'ai réalisé plusieurs analyses avec des TLM incluant différents degrés de complexités et en utilisant différents temps d'optimisation. Une synthèse de ces résultats est présentée dans la section A4. La section A1 décrit en détail le choix de l'état de référence du MRCC, pendant que la préparation des conditions initiales pour les VS est présentée dans la section A2. La section A3 décrit le choix du domaine et de la norme des VS.

A1. Définition des perturbations du MRCC et choix de l'état de référence

Dans les premières analyses, un ensemble de 10 (et ensuite 20) simulations MRCC réalisées avec la version 3.6.1 et pilotées avec les réanalyses NCEP, a été utilisé pour définir la VI pour juillet 1993 sur un domaine centré sur le Québec et l'ouest de l'Océan Atlantique (voir la figure A.1). La VI a été mesurée en utilisant la définition usuelle de la variance entre les membres de l'ensemble (X_m) par rapport à la moyenne d'ensemble ($\overline{X}^M$) :

$$\sigma_{IV}^2(x,y,z,t) = \frac{1}{M-1}\sum_{m=1}^{M}\left[X_m(x,y,z,t) - \overline{X}^M\right]^2 , \qquad (A.1)$$

et les perturbations MRCC ont été définies par rapport à la moyenne d'ensemble :

$$x'_m = X_m - \overline{X}^M . \qquad (A.2)$$

Ensuite, la VI a été mesurée en terme d'énergie totale moyenne des perturbations (équation 2.14).

Le modèle MRCC v3.6 utilisé pour effectuer cet ensemble diffère du modèle global utilisé à Dorval pour calculer les VS. En 2009, la version 5 du MRCC était stable et nous avons décidé d'utiliser cette version pour construire l'ensemble des simulations MRCC. Cette version du modèle est la variante climatique du modèle LAM utilisé à Dorval pour la prévision numérique du temps et elle utilise les mêmes modules dynamique et physique que le modèle global GEM. L'utilisation du MRCC5 au lieu de MRCCv3.6 nous a permis de rapprocher les simulations MRCC de la trajectoire du modèle GEM utilisé dans le calcul des VS. Un autre changement important dans la méthodologie vise l'état de référence utilisé pour calculer les perturbations MRCC et pour fournir les conditions initiales dans le calcul des VS. La moyenne d'ensemble, qui représentait l'état de référence dans les analyses antérieures, ne représente pas une solution des équations du modèle. Par conséquent, nous avons décidé de choisir comme état de référence une des simulations du MRCC5, qui elle est une solution du modèle, et nous avons opté pour le membre le plus proche du pilote en terme d'énergie totale intégrée sur le domaine MRCC5. Pour trouver le membre le plus proche du pilote, les étapes suivantes ont été parcourues :

- Interpolation des champs ERA-40 sur la grille du modèle MRCC5;
- Calcul des différences entre chaque simulation et le pilote;
- Calcul de l'énergie cinétique et de l'énergie potentielle pour les différences.

La figure A.2 présente l'évolution temporelle de ces différences en terme d'énergies cinétique et potentielle. Deux simulations se remarquent comme les plus proches des champs ERA-40 : la simulation NA11 et la simulation NA12. La moyenne temporelle sur le mois de décembre indique la simulation NA11 comme étant la plus proche et nous avons choisi cette simulation comme état de référence pour le calcul des simulations MRCC5.

A2. Préparation des conditions initiales pour le calcul des VS

Parce qu'on veut calculer les plus rapides perturbations à partir de l'état de référence MRCC, les conditions initiales pour les VS doivent être celles de l'état de référence MRCC. D'autre part, le calcul des VS est réalisé sur le domaine global et il nécessite des conditions initiales définies sur l'ensemble du globe. Pour que les conditions initiales satisfassent ces

deux conditions, j'ai combiné les données fournies par l'état de référence MRCC sur le domaine à aire limitée MRCC avec les données du pilote à l'extérieur du domaine MRCC.

Tel que précisé auparavant, dans les premières analyses, l'ensemble des simulations MRCC a été réalisé avec la version 3.6.1 du MRCC et piloté avec les réanalyses NCEP. Donc, il fallait combiner les champs MRCC définis sur une grille stéréographique polaire avec les champs de NCEP définis sur une grille latitude – longitude. Les prochaines étapes ont été réalisées :

- Sélection des champs MRCC sur les niveaux de pression identiques à ceux des champs NCEP (1000 hPa, 925 hPa, 850 hPa, 700 hPa, 600 hPa, 500 hPa, 400 hPa, 300 hPa, 250 hPa, 200 hPa, 150 hPa, 100 hPa, 70 hPa, 50 hPa, 30 hPa, 20 hPa, 10 hPa);
- Interpolation des champs MRCC de la grille stéréographique polaire à 45 km de résolution horizontale sur la grille latitude – longitude du NCEP à 2,5 dégrées résolution horizontale. Le résultat de l'interpolation produit des champs latitude – longitude avec les valeurs MRCC sur le domaine MRCC et des valeurs nulles sur le reste du globe.
- Réalisation d'un masque (grille latitude – longitude) qui permet de combiner les champs du MRCC avec ceux du NCEP en utilisant un passage linéaire entre les deux champs sur une distance de 9 points correspondant à l'éponge du MRCC sur la grille stéréographique polaire (voie la figure A.3). Le masque a la valeur 1 sur le domaine libre MRCC et 0 à l'extérieur du domaine MRCC. Un passage linéaire et réalisé dans la région de l'éponge.
- La combinaison des champs MRCC avec les champs du NCEP :
  CI=Masque * Champ MRCC + (1-Masque) * Champ NCEP.

Le passage graduel évite l'apparition des discontinuités dans le champ résultant. La même technique a été employée plus tard pour l'ensemble des simulations MRCC5 pilotées par ERA40.

A3. Le choix du domaine de la norme des VS

Initialement, les vecteurs VS ont été calculés en utilisant une norme initiale globale et une norme finale localisée à l'intérieur du domaine MRCC. Le désavantage principal de la norme initiale globale réside dans le fait qu'elle produit parfois des VS qui sont partialement à l'extérieur du domaine MRCC (voir comme exemple la figure 3.8). Ces perturbations ne peuvent pas se développer ou entrer à l'intérieur du domaine MRCC, les frontières du modèle étant rigides par construction et la VI du modèle nulle dans la région des frontières. Les seules perturbations qui peuvent croître et déterminer une augmentation de la VI, doivent être situées à l'intérieur du domaine MRCC. L'interpolation des VS globaux sur la grille MRCC peut éliminer la partie qui se trouve à l'extérieur du domaine au temps initial, mais ne permet pas d'identifier la partie correspondante pour les VS au temps final. Une norme au temps initial localisée à intérieur du domaine MRCC est nécessaire pour l'étude de la VI du MRCC. En conséquence, nous avons réalisé plusieurs tests qui utilisent des normes localisées aux temps initial ainsi qu'au temps final.

L'utilisation de normes initiale et finale localisées était un compromis imposé par l'impossibilité de calculer directement les VS sur le domaine MRCC. Cette solution n'est pas parfaite. Un des problèmes associés à la norme localisée réside dans le fait que les VS et les simulations MRCC utilisent des grilles différentes. Les VS sont sur une grille latitude – longitude non tournée, pendant que les simulations MRCC5 ont une grille tournée et les simulations MRCCv3.6 ont une grille stéréographique polaire. D'ici l'impossibilité de définir la norme des VS sur le domaine exact des simulations MRCC (voir par exemple les figures 3.5 et A.1). Par conséquent, il y a des régions MRCC où le calcul des VS au temps initial ne s'effectue pas. Un autre problème est causé par les VS au temps final. Malgré le fait que la norme au temps final est définie sur un domaine intérieur au domaine MRCC, les VS au temps final débordent la zone de la norme et donc ils sortent partiellement du domaine MRCC. Les VS calculés avec des normes localisées ne tiennent pas compte des frontières latérales du modèle.

En conclusion, le choix des VS globaux avec des normes localisées reste un compromis nécessaire, mais loin de la solution idéale des VS calculés directement sur un domaine à l'aire limitée.

A4. Application de la technique des VS pour un cas d'été

Tel que précisé auparavant, dans les premières analyses, l'ensemble de simulations MRCC a été réalisé avec la version 3.6.1 du MRCC et piloté avec les réanalyses NCEP. Le but était d'analyser la VI pour le mois de juillet 1993 sur un domaine de 120 par 120 points de grille à 45 km de résolution horizontale, centré sur le Québec et l'ouest de l'Océan Atlantique (voir la figure A.1). Les perturbations MRCC étaient calculées par rapport à la moyenne d'ensemble et les VS calculés en utilisant comme conditions initiales les champs du pilote : NCEP. La figure A.4 présente l'évolution temporelle de la VI du MRCCv3.6 pour juillet 1993. Nos analyses visaient la période de croissance associée au maximum principal du 20 juillet et surtout les projections des perturbations MRCC sur des VS au temps initial.

Très peu de ressemblances ont été trouvées entre les VS au temps initial et les perturbations MRCC au même moment. Un exemple est présenté dans les figures A.5 et A.6 où l'énergie totale des perturbations MRCC (en bleu pâle) est comparée avec celle de la partie projetée (en rouge) sur deux séries des VS au temps initial. Les deux séries comptent pour les 15 premiers VS calculés avec un TLM sec et avec la norme de l'énergie totale sèche localisée à l'intérieur du domaine MRCC au temps initial (contour rouge dans la figure A.1) ainsi qu'au temps final. La première série (figure A.5) utilise les conditions initiales du NCEP pour 16 juillet 1993 à 00Z, pendant que la deuxième série (figure A.6) utilise les conditions initiales du NCEP pour 18 juillet 1993 à 00Z. Les deux séries ont un temps d'optimisation de 48 heures. La projection au temps initial est approximativement 1 % pour la série no. 1 et de 5 % pour la série no. 2. Cependant, la projection sur les VS au temps final présente des valeurs beaucoup plus grandes : jusqu'à 29 % pour les VS de la série no. 1 et 35 % pour les VS de la série no. 2.

Dans ces études, nous nous sommes intéressés surtout à la projection des perturbations MRCC sur les VS au temps initial. Les résultats obtenus montraient des VS au temps initial très différents des perturbations MRCC. Ils avaient un caractère fortement barocline et ils s'étendaient sur tout le domaine, contrairement aux perturbations MRCC qui étaient localisées sur une petite région et présentaient un caractère plutôt barotrope. De plus, la région où les VS étaient situés au temps initial ne correspondait pas au maximum de VI dans

MRCC. Par contre, aux temps final, les VS avaient évolué vers des structures semblables aux perturbations MRCC et une bonne concordance a été observée dans la position des maximums des énergies totales intégrées sur la verticale.

À ce moment, notre explication pour le pourcentage très petit obtenu pour la projection au temps initial et son augmentation au temps final était la suivante :

- Au temps initial, les sources des instabilités développées dans MRCC avaient un caractère de très petite échelle et elles étaient submergées dans une masse d'autres perturbations de grande échelle, qui eux ont déjà évolué; d'ici l'impossibilité de les détecter visuellement.
- Dans l'intervalle de temps suivant, ces petites sources d'instabilité sont développées rapidement et au temps final ils expliquaient approximativement 30 % de la VI.

Pour valider cette hypothèse, nous avons réalisé un ensemble de simulations MRCC (ensemble MRCC « B ») avec les conditions initiales du NCEP du 18 juillet 1993 perturbées avec la partie du premier ensemble MRCC (ensemble MRCC « A ») projetée sur les VS de la série no. 2 au temps initial. La figure A.7 présente l'évolution temporelle de l'énergie totale correspondant à l'ensemble « A » (en jaune) et à l'ensemble « B » (en noir). Parce que les perturbations des conditions initiales de l'ensemble « B » sont très petites, il y a une grande différence entre la VI du premier ensemble et celle du deuxième ensemble au moment initial. Cependant, dans les deux cas on obtient des variations temporelles semblables, les positions des maximums et minimums étant approximativement les mêmes. Le premier maximum est atteint un pas de temps plus tard que dans l'ensemble « A » et représente seulement 29 % de la valeur de l'ensemble « A ». Dans la figure A.8, l'énergie totale de l'ensemble « A » et de sa partie projetée sur les VS de la série no 2 au temps final sont comparées avec l'énergie totale de l'ensemble « B ». Les perturbations des conditions initiales de l'ensemble « B » (les VS de la série no 2 au temps initial) évoluent dans MRCC et finissent par expliquer approximativement la même portion de la VI que les VS au temps final. La figure A.7 montre que vers la fin du mois, les deux ensembles tendent à avoir la même VI.

La grande différence du moment initial et cette convergence vers la même valeur vers la fin du mois peuvent être expliquées par le fait que les VS au temps initial ne sont pas sur

l'attracteur du système; ils représentent des perturbations qui n'existent pas dans les simulations MRCC. D'où le pourcentage très petit obtenu pour toutes les séries de VS au temps initial. Vers le temps final, les VS tournent rapidement vers l'attracteur et ils finissent par présenter une bonne partie des perturbations existantes dans les simulations MRCC.

Pour approcher les VS au temps initial des perturbations MRCC, plusieurs tests ont été réalisés avec des séries de VS calculés avec différents temps d'optimisations (24h, 36h et 48h), à partir des conditions initiales fournies par NCEP ou par la moyenne d'ensemble MRCCv3.6, et avec des normes initiales et finales localisées sur des portions du domaine MRCC centrées sur le maximum de la VI. En général, les pourcentages augmentent légèrement au temps initial si les conditions initiales sont fournies par la moyenne d'ensemble. Par exemple la projection des perturbations MRCCv3.6 sur les VS calculés avec un OTI de 24 h, la norme initiale de l'énergie totale sèche localisée sur le domaine MRCC (contour magenta dans la figure A.1) et avec les conditions initiales de la moyenne d'ensemble du 17 juillet 1993 varie entre 4 % et 13 % au temps initial (série des SV no 3 représentée en magenta dans la figure A.9a) et entre 8 % et 35 % au temps final (série des SV no 3 représentée en magenta dans la figure A.9b). En diminuant fortement le domaine de la norme initiale et d'analyse, les VS sont empêchés d'avoir une structure barocline. Cet artifice a augmenté la projection au temps initial pour des valeurs entre 2 % et 42 % pour les VS avec le plus petit domaine, un OTI de 24h et conditions initiales de la moyenne d'ensemble MRCC (série des SV représentée en bleu dans la figure A.9a); la projection sur les VS correspondent au temps final variait entre 12 % et 52 % (série des SV représentée en bleu dans la figure A.9a). Le domaine de la norme au temps initial pour ce cas est présenté dans la figure A.1 par le contour vert. L'analyse dans ce cas est restreinte seulement à la petite région de la norme et fait abstraction des perturbations MRCC situées à l'extérieur de ce domaine. L'utilisation d'un très petit domaine force les VS à avoir une structure verticale proche de celle des perturbations MRCC au temps initial; cependant l'énergie totale au temps initial continue à être dominée par la composante potentielle contrairement au cas des perturbations MRCC.

D'autres tests ont été réalisés aussi en utilisant le TLM humide au lieu du TLM sec et la norme de l'énergie totale humide au lieu de la norme de l'énergie totale sèche pour

différentes conditions initiales (16 juillet à 12Z, 17 juillet à 00Z, 18 juillet à 00Z) et différents temps d'optimisations (48h, 36h, 24h, 12h). L'utilisation d'un TLM humide a résulté dans l'apparition de nouvelles structures dans les séries des VS qui n'étaient pas présentes dans le cas d'un TLM sec. Malgré l'importance évidente de l'inclusion des processus humides dans l'analyse, la projection des perturbations MRCC sur les VS au temps final était tombée à 8 % pour le cas des VS calculés à partir des conditions initiales NCEP et avec une norme de l'énergie totale humide couvrant le domaine MRCC (contour rouge dans la figure A.1). Le problème provenait de la définition de la norme humide qui employait une pondération relativement arbitraire pour le terme de l'humidité (équation 1.25) réalisée par l'intermédiaire du coefficient ε. Le résultat était une énergie totale de perturbations MRCC dominée par le terme d'humidité tout au long de la simulation. Les VS calculés avec la norme de l'énergie totale humide étaient dominés par l'énergie potentielle au temps initial et par l'énergie humide au temps final.

Par conséquent, nous avons renoncé à la norme humide dans le calcul de la VI ainsi que dans le calcul des SV et nous avons calculé les VS avec des normes initiales et finales définies par l'énergie totale sèche, mais avec un TLM humide. Dans ce cas, le pourcentage de la projection ne change pas visiblement par rapport au cas où le TLM sec est utilisé. Un exemple est présenté dans les figures A.9a et A.9b par la série de VS no 3 représentée en vert. La seule différence entre la série no 2 (en bleu) et la série no 3 (en vert) est l'utilisation d'un TLM humide pour la série no 3 et un TLM sec pour la série no 2. Les projections présentent des valeurs semblables au temps initial ainsi qu'au temps final.

La présence d'importantes perturbations dans le champ d'humidité MRCC, ainsi que la présence de nouvelles structures dans les séries des VS obtenues avec le TLM humide, qui n'étaient pas présentes dans le cas d'un TLM sec, confirme la nécessité de l'utilisation d'un TLM humide. Cependant, nos résultats ne présentaient aucun changement majeur et suggéraient une révision de la représentation des processus humides dans le TLM. En conséquence, nous avons décidé de concentrer l'analyse de la VI sur un cas d'hiver sur le continent de l'Amérique du Nord. Dans ce cas, l'humidité ne jouait pas un rôle important dans la VI et les perturbations MRCC présentaient des caractéristiques baroclines comme les VS produits avec le TLM sec. Tel que précisé auparavant, à ce moment, la version 5 du

MRCC était disponible sur les serveurs à l'UQAM et nous avons utilisé cette version du modèle pour construire l'ensemble des simulations pour décembre 1992, ce qui nous a permis de diminuer les différences entre le modèle à l'aire limitée et le TLM.

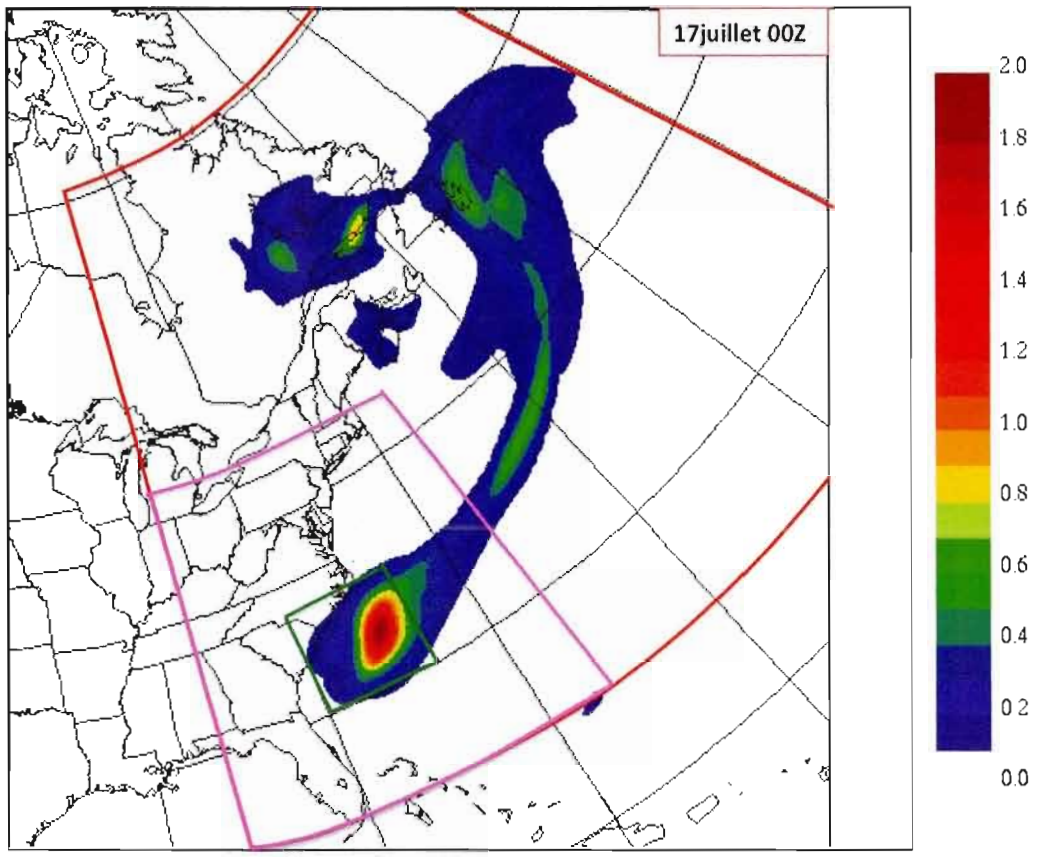


Figure A.1 : Distribution horizontale de l'énergie totale de l'ensemble MRCCv3.6, le 16 juillet 1993 à 12Z. L'état de référence est la moyenne d'ensemble des 10 simulations MRCC. Le contour vert (rouge) indique le plus petit (grand) domaine utilisé pour la norme localisée initiale pour les séries des VS. Le contour magenta correspond à une norme avec un domaine intermédiaire.

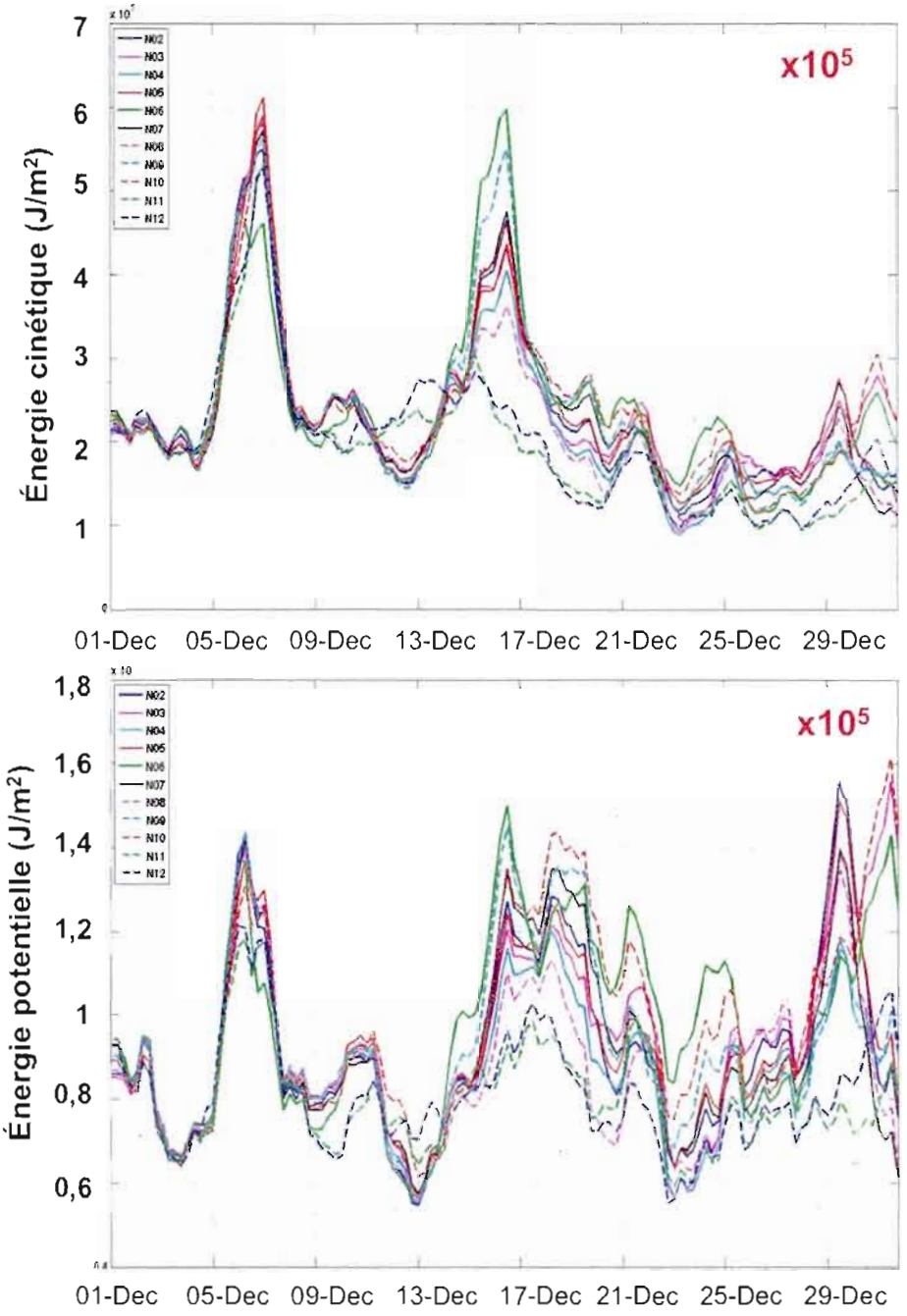


Figure A.2 : Evolution temporelle des différences simulations MRCC – ERA-40. Les différences sont exprimées en termes d'énergie cinétique et d'énergie potentielle.

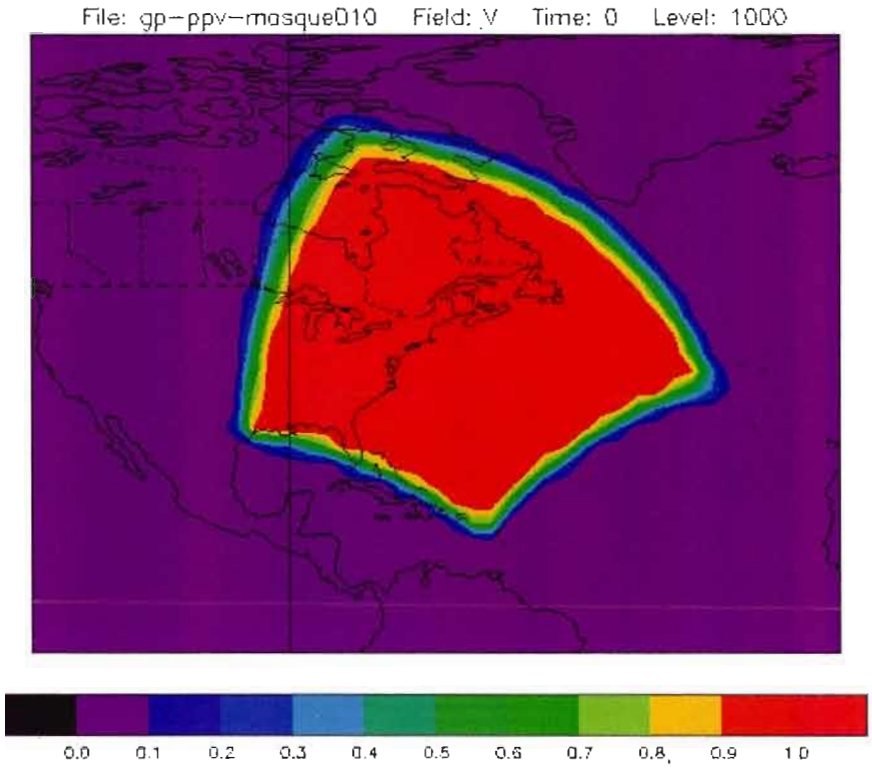


Figure A.3 : Masque qui permet de combiner les champs du MRCC avec les champs du NCEP.

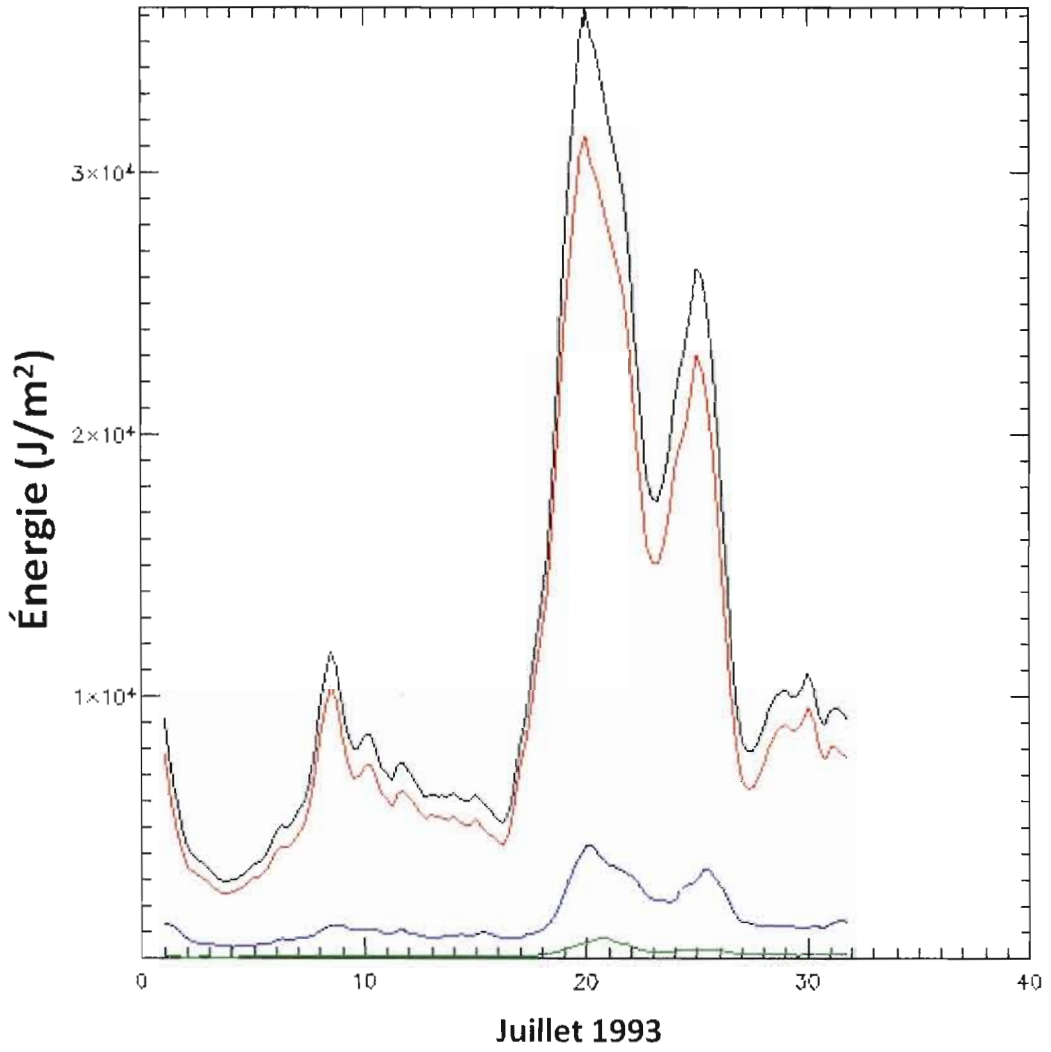
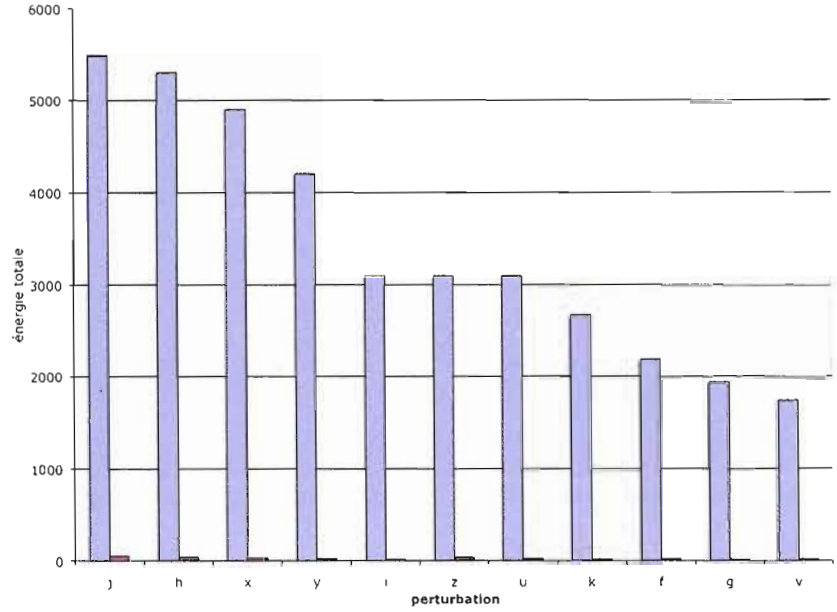


Figure A.4 : Évolution temporelle de la VI du MRCCv3.6 pour juillet 1993. L 'énergie totale est représentée en noir, l'énergie cinétique en rouge, l'énergie potentielle en bleu et le terme de la pression en surface en vert. Tous les termes sont exprimés en Jm^{-2} et sont le résultat de l'intégration sur le domaine horizontale MRCC et sur la verticale entre les niveaux de 100 hPa et 1000 hPa.

a) VS série no 1; temps initial (16 juillet 1993)



b) VS série no 1; temps final (18 juillet 1993)

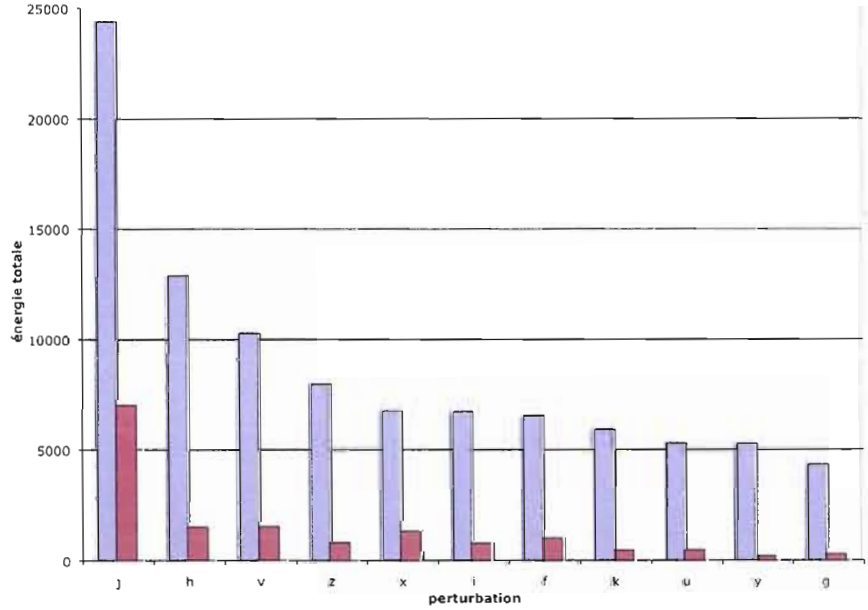
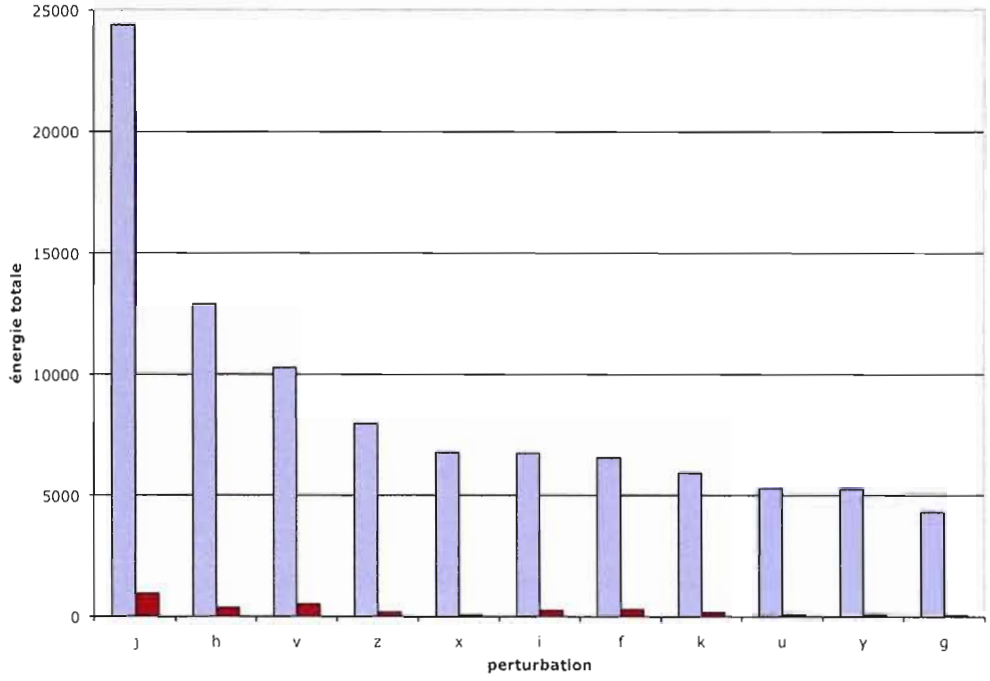


Figure A.5 : Énergie totale des perturbations MRCCv3.6 (en bleu) et des parties projetées (en rouge) sur la série des VS no 1 au temps initial (a) et au temps final (c). Le domaine de la norme au temps initial correspond au contour rouge dans la figure A.1.

a) VS série no 2; temps initial (18 juillet 1993)



b) VS série no 2; temps final (20 juillet 1993)

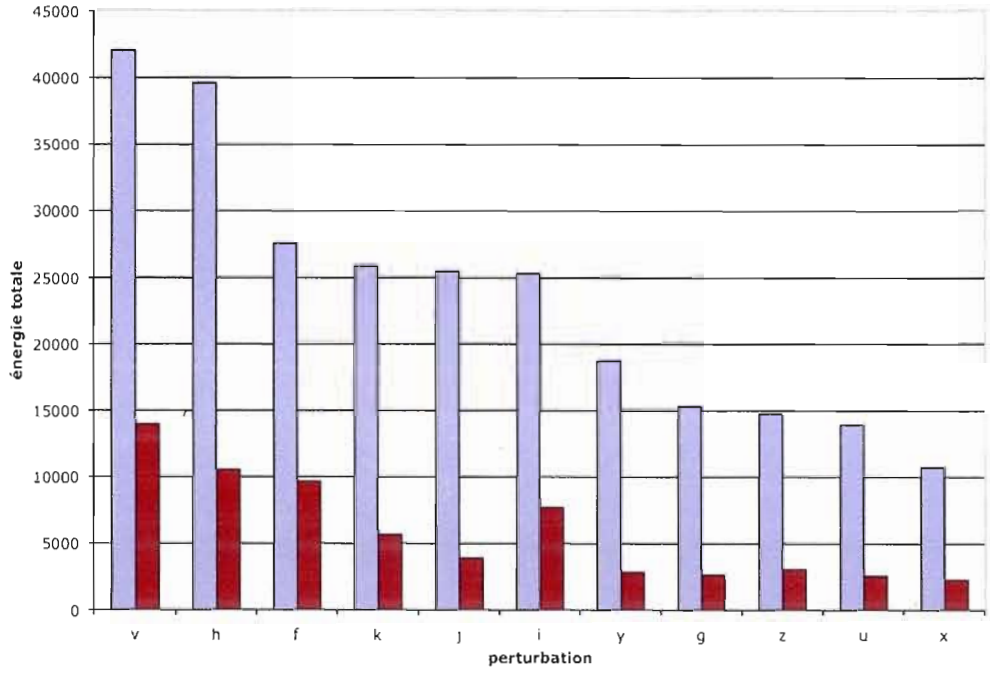


Figure A.6 : Énergie totale des perturbations MRCCv3.6 (en bleu) et des parties projetées (en rouge) sur la série des VS no 2 au temps initial (a) et au temps final (c). Le domaine de la norme au temps initial correspond au contour rouge dans la figure A.1.

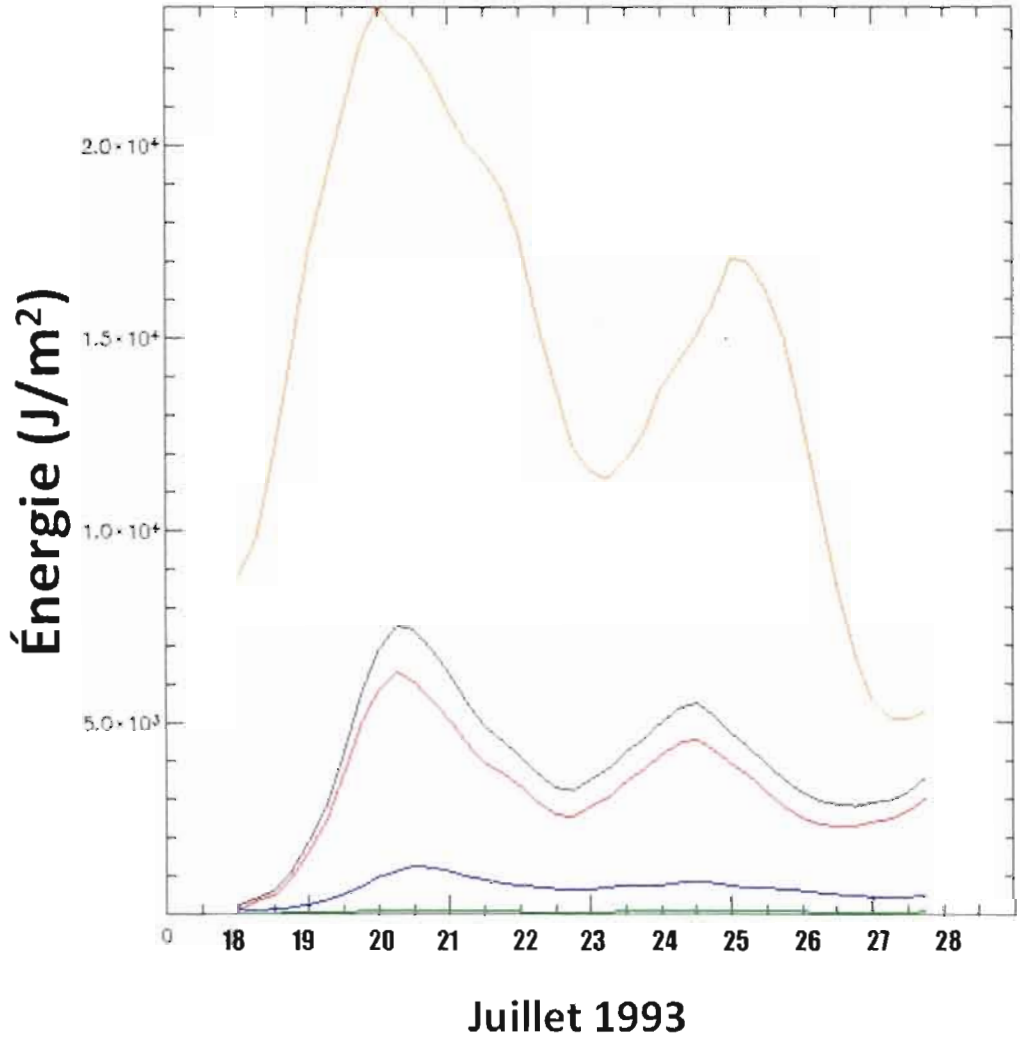


Figure A.7 : Évolution temporelle de la VI de l'ensemble MRCC « A » (en jaune – énergie totale) et ensemble MRCC « B » (l'énergie cinétique en rouge, l'énergie potentielle en bleu, le terme de la pression en surface en vert et l'énergie totale en noir). Tous les termes sont exprimés en Jm^{-2} et ils sont le résultat de l'intégration sur le domaine horizontale MRCC et sur la verticale entre les niveaux de 100 hPa et 1000 hPa.

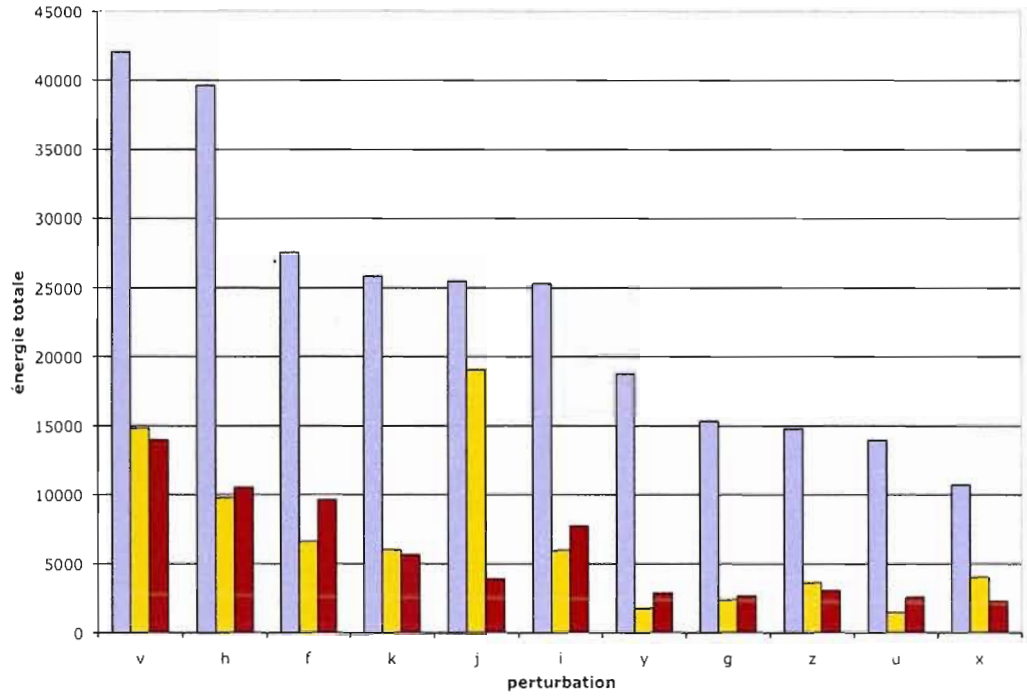


Figure A.8 : Énergie totale des perturbations de l'ensemble MRCC « A » (en bleu) et de sa partie projetée sur la série de VS no. 2 au temps final (en rouge) et énergie totale des perturbations de l'ensemble MRCC « B ».

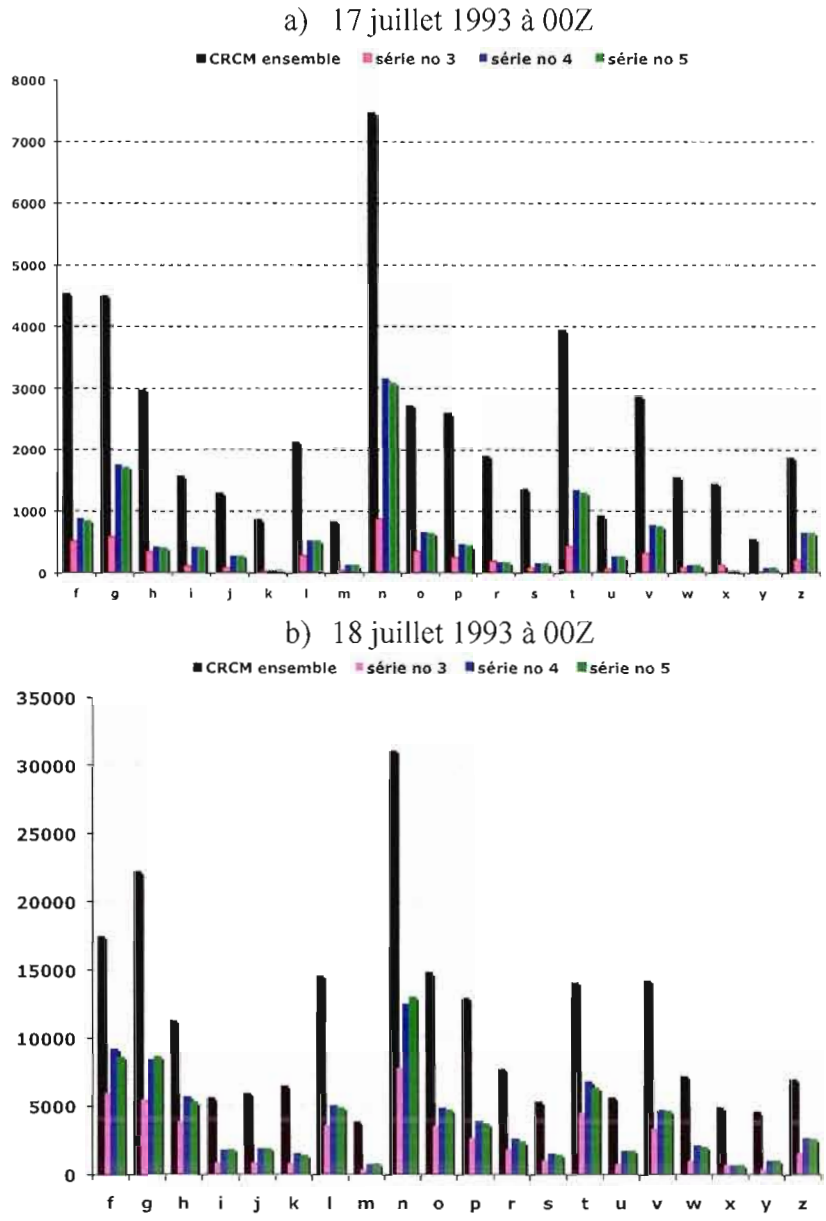


Figure A.9 : Énergie totale des perturbations MRCCv3.6 (en bleu pâle) et des parties projetées sur les séries des VS no 3 (en magenta), no 4 (en bleu foncé) et no 5 (en vert) au temps initial (a) et au temps final (b). Le domaine de la norme au temps initial correspond dans la figure A.1 au contour magenta pour la série no 3 et au contour vert pour les séries no 4 et 5.

RÉFÉRENCES

Alexandru, A., R. de Elía et R. Laprise. 2007. «Internal variability in Regional Climate Downscaling at the seasonal scale ». *Monthly Weather Review*, vol. 135, no 9, p. 3221–3238.

_____, R. de Elía, R. Laprise, L. Šeparović et S. Biner. 2009. « Sensitivity Study of Regional Climate Model Simulations to Large-Scale Nudging Parameters ». *Monthly Weather Review*, vol. 137, no 5, p. 1666-1686.

Badger, J., et B. J. Hoskins. 2001. « Simple Initial Value Problems and Mechanisms for Baroclinic Growth ». *Journal of Atmospheric Science*, vol. 58, p. 38-49.

Barkmeijer, J. 1996. « Constructing fast-growing perturbation for the nonlinear regime ». *Journal of Atmospheric Science*, vol. 53, p. 2838-2851.

_____, R. Buizza et T. N. Palmer. 1999. « 3D-Var Hessian singular vectors and their potential use in the ECMWF Ensemble Prediction System ». *Quarterly Journal of the Royal Meteorological Society*, vol. 125, p. 2333-2351.

_____, R. Buizza, T. N. Palmer, K. Puri et Mahfouf J. 2001. « Tropical singular vectors computed with linearized diabatic physics ». *Quarterly Journal of the Royal Meteorological Society*, vol. 127, p. 685-708.

_____, T. Iversen et T. N. Palmer. 2003. « Forcing singular vectors and other sensitive model structures ». *Quarterly Journal of the Royal Meteorological Society*, vol. 129, p. 2401-2423.

Bélair, S., L.-P. Crevier, J. Mailhot, B. Bilodeau et Y. Delage. 2003. « Operational implementation of the ISBA land surface scheme in the Canadian regional weather forecast model. Part I: warm season results ». *Journal of Hydrometeorology*, vol. 4, p. 352–370.

_____, J. Mailhot, C. Girard et P. Vaillancourt. 2005. « Boundary layer and shallow cumulus clouds in a medium-range forecast of a large-scale weather system ». *Monthly Weather Review*, vol. 133, p. 1938–1959.

Borges, M. D., et D. L. Hartmann. 1992. « Barotropic instability and optimal perturbations of observed nonzonal flows ». *Journal of Atmospheric Science*, vol. 49, p. 335-353.

Buehner, M., et A. Zadra. 2006. « Impact of flow-dependent analysis-error covariance norms on extratropical singular vectors ». *Quarterly Journal of the Royal Meteorological Society*, vol. 132, p. 625-646.

_____, et A. Mahidjiba. 2010. « Sensitivity of Global Ensemble Forecasts to the Initial Ensemble Mean and Perturbations: Comparison of EnKF, Singular Vector and 4D-Var Approaches ». *Monthly Weather Review*, vol. 138, p. 3886-3904.

Buizza, R. 1994. « Sensitivity of Optimal Unstable Structures ». *Quarterly Journal of the Royal Meteorological Society*, vol. 120, p. 429-451.

_____. 1998. « Impact of horizontal diffusion on T21, T42, and T63 singular vectors ». *Journal of the Atmospheric Sciences*, vol. 55, p. 1069-1083.

_____. 2010. « The Value of a Variable Resolution Approach to Numerical Weather Prediction ». *Monthly Weather Review*, vol. 138, p. 1026-1042.

_____, J.-R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt et F. Vitart. 2007a. « The new ECMWF VAREPS ». *Quarterly Journal of the Royal Meteorological Society*, vol. 133, p. 681-695.

_____, C. Cardinali, G. Kelly et J.-N. Thépaut. 2007b. « The value of observations. II: The value of observations located in singular-vector-based target areas ». *Quarterly Journal of the Royal Meteorological Society*, vol. 133, p. 1817-1832.

_____, R. Gelaro, F. Molteni et T. N. Palmer. 1997. « The impact of increased resolution on predictability studies with singular vectors ». *Quarterly Journal of the Royal Meteorological Society*, vol. 123, p. 1007-1033.

_____, M. Leutbecher et L. Isaksen. 2008. « Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System ». *Quarterly Journal of the Royal Meteorological Society*, vol. 134, p. 2051-2066.

_____, et A. Montani. 1999. « Targeting observations using singular vectors ». *Journal of the Atmospheric Sciences*, vol. 56, p. 2965-2985.

_____, et T. N. Palmer. 1995. « The singular-vector structure of the atmospheric global circulation ». *Journal of the Atmospheric Sciences*, vol. 52, p. 1434-1456.

_____, J. Tribbia, F. Molteni et T. N. Palmer. 1993. « Computation of optimal unstable structures for a numerical weather prediction model ». *Tellus*, vol. 45A, p. 388-407.

Cardinali, C., R. Buizza, G. Kelly, M. Shapiro et J.-N. Thépaut. 2007. « The value of observations - Part III: Influence of weather regimes on targeting ». *Quarterly Journal of the Royal Meteorological Society*, vol. 133, p. 1833-1842.

Caya, D., et S. Biner. 2004. « Internal Variability of RCM Simulations over an Annual Cycle ». *Climate Dynamics*, vol. 22, p. 33-46.

Chen, J.-H., M. S. Peng, C. A. Reynolds et C.-C. Wu. 2009. « Interpretation of Tropical Cyclone Forecast Sensitivity from the Singular Vector Perspective ». *Journal of the Atmospheric Sciences*, vol. 66, p. 3383-3400.

Cheng, Y., Y. Tang, X. Zhou, P. Jackson et D. Chen. 2010. « Further analysis of singular vector and ENSO predictability in the Lamont model—Part I: Singular vector and the control factors ». *Climate Dynamics*, vol. 35, p. 807-826.

Christensen, O. B., M. A. Gaertner, J. A. Prego et J. Polcher. 2001. « Internal variability of regional climate models ». *Climate Dynamics*, vol. 17, p. 875-887.

Coutinho, M. M., B. J. Hoskins et R. Buizza. 2004. « The influence of physical processes on extratropical singular vectors ». *Journal of the Atmospheric Sciences*, vol. 61, p. 195-209.

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch et A. Staniforth. 1998. « The operational CMC-MRB global environmental multiscale (GEM) model. Part I: Design considerations and formulation ». *Monthly Weather Review*, vol. 126, p. 1373-1395.

Davies, H. C. 1976. « A lateral boundary formulation for multi-level prediction models ». *Quarterly Journal of the Royal Meteorological Society*, vol. 102, p. 405-418.

Descamps, L., D. Ricard, A. Joly et P. Arbogast. 2007. « Is a Real Cyclogenesis Case Explained by Generalized Linear Baroclinic Instability? » *Journal of the Atmospheric Sciences*, vol. 64, p. 4287-4308.

Descamps, L., et O. Talagrand. 2007. « On some aspects of the definition of initial conditions for ensemble prediction ». *Monthly Weather Review*, vol. 135, p. 3260-3272.

Diaconescu, E. P., R. Laprise, et A. Zadra. 2011. « Singular Vector Decomposition of the Internal Variability of the Canadian Regional Climate Model ». Submitted to *Climate Dynamics*.

Ehrendorfer, M., et R. M. Errico. 1995. « Mesoscale Predictability and the Spectrum of Optimal Perturbations ». *Journal of the Atmospheric Sciences*, vol. 52, p. 3475-3500.

_____, R. M. Errico et K. D. Raeder. 1999. « Singular vector perturbation growth in a primitive equation model with moist physics ». *Journal of the Atmospheric Sciences*, vol. 56, p. 1627-1648.

_____, et J. J. Tribbia. 1997. « Optimal prediction of forecast error covariances through singular vectors ». *Journal of the Atmospheric Sciences*, vol. 53, p. 286-313.

Errico, R. M., M. Ehrendorfer et K. D. Raeder. 2001. « The spectra of singular values in a regional model ». *Tellus*, vol. 53A, p. 317-332.

_____, et K. D. Raeder. 1999. « An examination of the accuracy of the linearization of a mesoscale model with moist physics ». *Quarterly Journal of the Royal Meteorological Society*, vol. 125, p. 169-195.

Farrell, B. F. 1982. « The initial growth of disturbances in a baroclinic flow ». *Journal of the Atmospheric Sciences*, vol. 39, p. 1663-1686.

_____. 1985. « Transient growth of damped baroclinic waves ». *Journal of the Atmospheric Sciences*, vol. 42, p. 2718-2727.

_____. 1989. « Optimal excitation of baroclinic waves ». *Journal of the Atmospheric Sciences,* vol. 46, p. 1193-1206.

_____ et P. J. Ioannou. 1996. « Generalized stability. Part I: Autonomous operators ». *Journal of the Atmospheric Sciences,* vol. 53, p. 2025–2040.

Frederiksen, J. S. 2000: « Singular Vectors, Finite-Time Normal Modes, and Error Growth during Blocking ». *Journal of the Atmospheric Sciences*, vol. 57, p. 312-333.

Gauthier, P., M. Tanguay, S. Laroche, S. Pellerin, et J. Morneau. 2007. « Extension of 3D-Var to 4D-Var: Implementation of 4D-Var at the Meteorological Service of Canada ». *Monthly Weather Review*, vol. 135, p. 2339-2354.

Gelaro, R., R. Buizza, T. N. Palmer et E. Klinker. 1998. « Sensitivity Analysis of Forecast Errors and the Construction of Optimal Perturbations Using Singular Vectors ». *Journal of the Atmospheric Sciences*, vol. 55, p. 1012-1037.

Gilmour, I., L. A. Smith et R. Buizza. 2001. « Linear Regime Duration: Is 24 Hours a Long Time in Synoptic Weather Forecasting? ». *Journal of the Atmospheric Sciences*, vol. 58, p. 3525-3539.

Giorgi, F., et X. Bi. 2000. « A study of IV of regional climate model ». *Journal of Geophysical Research*, vol. 105, p. 29503- 29521.

Gleckler, P. 1996. « AMIP II guidelines ». AMIP Newsletter, No. 8, PCMDI/LLNL. http://www-pcmdi.llnl.gov/projects/amip/NEWS.

Hakim, G. J. 2000. « Role of nonmodal growth and nonlinearity in cyclogenesis initial-value problems ». *Journal of the Atmospheric Sciences*, vol. 57, p. 2951-2967.

Harnisch, F., et M. Weissmann. 2010. « Sensitivity of typhoon forecasts to different subsets of targeted dropsonde observations ». *Monthly Weather Review*, vol. 138, p. 2664-2680.

Hartmann, D. L., R. Buizza et T. N. Palmer. 1995. « Singular vectors: The effect of spatial scale on linear growth of disturbances ». *Journal of the Atmospheric Sciences*, vol. 52, p. 3885-3894.

Hoskins, B. J., R. Buizza et J. Badger. 2000. « The nature of singular growth and structure ». *Quarterly Journal of the Royal Meteorological Society*, vol. 126, p. 1565–1580.

_____, et M. M. Coutinho. 2005. « Moist singular vectors and the predictability of some high impact European cyclones ». *Quarterly Journal of the Royal Meteorological Society*, vol. 131, p. 581-601.

Houtekamer, P. L. 1995. « The construction of optimal perturbations ». *Monthly Weather Review*,

vol. 123, p. 2888–2898.

_____, et L. M. Herschel. 2005. « Ensemble Kalman filtering ». *Quarterly Journal of the Royal Meteorological Society*, vol. 131, p. 3269-3289.

Kain, J. S., et J. M. Fritsch. 1990. « A one-dimensional entraining/detraining plume model and application in convective parameterization ». *Journal of Atmospheric Science*, vol. 47, p. 2784–2802.

Kalnay, E. 2002. « Atmospheric modeling, data assimilation and predictability ». Cambridge University Press, 341 p.

Kelly, G., J. N. Thépaut, R. Buizza et C. Cardinali. 2007. « The value of observations - Part I: Data denial experiments for the Atlantic and the Pacific ». *Quarterly Journal of the Royal Meteorological Society*, vol. 133, p. 1803-1815.

Kim, H. M., et B.-J. Jung. 2009a. « Influence of Moist Physics and Norms on Singular Vectors for a Tropical Cyclone ». *Monthly Weather Review*, vol. 137, p. 525-543.

_____, et B.-J. Jung. 2009b. « Singular Vector Structure and Evolution of a Recurving Tropical Cyclone ». *Monthly Weather Review*, vol. 137, p. 505-524.

Kleeman, R., Y. Tang et A. M. Moore. 2003. « The calculation of climatically relevant singular vectors in the presence of weather noise as applied to the ENSO problem ». *Journal of the Atmospheric Sciences*, vol. 60, p. 2856-2868.

Komori, T., et T. Kadowaki. 2010. « Resolution Dependence of Singular Vectors Computed for Typhoon SINLAKU ». *SOLA Scientific Online Letters on the Atmosphere*, vol. 6, p. 045-048.

Kuo, H. L. 1965. « On formation and intensification of tropical cyclones through latent heat release by cumulus convection ». *Journal of Atmospheric Science*, vol. 22, p. 4063.

Lacarra, J., et O. Talagrand. 1988. « Short-range evolution of small perturbations in a barotropic model ». *Tellus*, vol. 40A, p. 81-95.

Langland, R.H. 2005. « Issues in targeted observing ». *Quarterly Journal of the Royal Meteorological Society*, vol. 131. p. 3409-3425.

Laprise, R. 1992. « The Euler equations of Motion with Hydrostatic Pressure as an Independent Variable ». *Monthly Weather Review*, vol. 130, p. 197-207.

_____, R. de Elía, D. Caya, S. Biner, P. Lucas-Picher, E. P. Diaconescu, M. Leduc, A. Alexandru et L. Šeparović. 2008. « Challenging some tenets of Regional Climate Modelling ». *Meteorology and Atmospheric Physics, Special Issue on Regional Climate Studies*, vol. 100, p. 3–22.

Leutbecher, M., 2007. « On the representation of initial uncertainties with multiple sets of singular vectors optimized for different criteria ». *Quarterly Journal of the Royal Meteorological Society*, vol. 133, p. 2045-2056.

Li, J., et H. W. Barker. 2005. « A radiation algorithm with correlated-k distribution. Part I: local thermal equilibrium ». *Journal of Atmospheric Science*, vol. 62, p. 286–309.

Lorenz, E. N. 1963. « Deterministic non-periodic flow ». *Journal of Atmospheric Science*, vol. 20, p. 130-141.

_____. 1965. « A study of the predictability of a 28-variable atmospheric model ». *Tellus*, vol. 17A, p. 321-333.

Lucas-Picher, P., D. Caya, R. de Elia et R. Laprise. 2008a. « Investigation of regional climate models' internal variability with a ten-member ensemble of ten-year simulations over a large domain ». *Climate Dynamics*, vol. 31, p. 927-940.

_____, D. Caya, S. Biner et R. Laprise. 2008b. « Quantification of the lateral boundary forcing of a regional climate model using an ageing tracer ». *Monthly Weather Review*, vol. 136, 4980-4996.

Majumdar, S. J., S. D. Aberson, H. C. Bishop, R. Buizza, M. S. Peng et C. A. Reynolds. 2006. « A Comparison of Adaptive Observing Guidance for Atlantic Tropical Cyclones ». *Monthly Weather Review*, vol. 134, p. 2354-2372.

McFarlane, N. A. 1987. « The effect of orographically excited gravity-wave drag on the circulation of the lower stratosphere and troposphere ». *Journal of Atmospheric Science*, vol. 44, p. 1175–1800.

Molteni, F., R. Buizza, T. N. Palmer et T. Petroliagis. 1996. « The new ECMWF ensemble prediction system: Methodology and validation ». *Quarterly Journal of the Royal Meteorological Society*, vol. 122, p. 73-119.

_____, et T. N. Palmer. 1993. « Predictability and finite-time instability of the northern winter circulation ». *Quarterly Journal of the Royal Meteorological Society*, vol. 119, p. 269-298.

Montani, A., et A. J. Thorpe. 2002. « Mechanisms leading to singular-vector growth for FASTEX cyclones ». *Quarterly Journal of the Royal Meteorological Society*, vol. 128, p. 131-148.

Nikiema, O., et R. Laprise. 2010. « Diagnostic budget study of the internal variability in ensemble simulations of the Canadian RCM ». *Climate Dynamics*, vol. 36, p. 2313-2337.

Palmer, T. N., R. Buizza, M. Leutbecher, R. Hagedorn, T. Jung, M. Rodwell, F. Vitart, J. Berner, E. Hagel, A. Lawrence, F. Pappenberger, Y.-Y. Park, L. van Bremen, I. Gilmour et L. Smith. 2007. « The ECMWF Ensemble Prediction System: Recent and on-going developments ». In *A paper presented at the 36th Session of the ECMWF Scientific Advisory Committee.* ECMWF Research Department Technical Memorandum No. 540, available from ECMWF,

Shinfield Park, Reading RG2-9AX.

_____, R. Buizza, F. Molteni, Y.-Q. Chen et S. Corti. 1994. « Singular vectors and the predictability of weather and climate ». *Philosophical Transactions of the Royal Society*, vol. 348, p. 459-475.

_____, R. Gelaro, J. Barkmeijer et R. Buizza. 1998. « Singular vectors, metrics and adaptive observations ». *Journal of the Atmospheric Sciences*, vol. 55, p. 633-653.

Pellerin, G., L. Lefaivre, P. Houtekamer et C. Girard. 2003. « Increasing the horizontal resolution of ensemble forecasts at CMC ». *Nonlinear Processes in Geophysics*, vol. 10, p. 463-468.

Puri, K., J. Barkmeijer et T. N. Palmer. 2001. « Ensemble prediction of tropical cyclones using targeted diabatic singular vectors ». *Quarterly Journal of the Royal Meteorological Society*, vol. 127, p. 709-731.

Rabier, F., P. Gauthier, C. Cardinali, R. Langland, M. Tsyrulnikov, A. C. Lorenc, R. Gelaro, P. Steinle et K. Koizumi. 2008. « An update on THORPEX-related research in Data Assimilation and Observing Strategies ». *Nonlinear Processes in Geophysics*, vol. 15, p. 1-14.

Rapaić, M., M. Leduc et R. Laprise. 2010. « Evaluation of the internal variability and estimation of the downscaling ability of the canadian regional climate model for different domain sizes over the north atlantic region using the big-brother experimental approach ». *Climate Dynamics*, vol. 36, p. 1979-2001.

Reynolds, C. A., et R. M. Errico. 1999. « Convergence of Singular Vectors toward Lyapunov Vectors ». *Monthly Weather Review*, vol. 127, p. 2309-2323.

_____, M. S. Peng et J. H. Chen. 2009. « Recurving Tropical Cyclones: Singular Vector Sensitivity and Downstream Impacts ». *Monthly Weather Review*, vol. 137, p. 1320-1337.

_____, M. S. Peng, S. J. Majumdar, S. D. Aberson, H. C. Bishop et R. Buizza. 2007. « Interpretation of Adaptive Observing Guidance for Atlantic Tropical Cyclones ». *Monthly Weather Review*, vol. 135, p. 4006-4029.

_____, et T. E. Rosmond. 2003. « Nonlinear growth of singular vector-based perturbations ». *Quarterly Journal of the Royal Meteorological Society*, vol. 129, p. 3059-3078.

Rinke, A., et K. Dethloff. 2000. « On the sensitivity of a regional Arctic climate model to initial and boundary conditions ». *Climate Research*, vol. 14, p. 101-113.

_____, P. Marbaix et K. Dethloff. 2004. « Internal Variability in Arctic regional climate simulations: case study for the Sheba year ». *Climate Research*, vol. 27, p. 197-209.

Šeparović, L., R. de Elía et R. Laprise. 2008. « Reproducible and Irreproducible Components in Ensemble Simulations with a Regional Climate Model ». *Monthly Weather Review*, vol. 136, p. 4942-4961.

Snyder, C., et G. J. Hakim. 2005. « Cyclogenetic perturbations and analysis errors decomposed into singular vectors ». *Journal of Atmospheric Science*, vol. 62, p. 2234-2247.

Sundqvist, H., E. Berge et J. E. Kristjansson. 1989. « Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model ». *Monthly Weather Review*, vol. 117, p. 1641–1657.

Szunyogh, I., E. Kalnay et Z. Toth. 1997. « A comparison of Lyapunov and optimal vectors in a low-resolution GCM ». *Tellus*, vol. 49A, p. 200-227.

Tanguay, M., P. Bartello et P. Gauthier. 1995. « Four-dimensional data assimilation with a wide range of scales ». *Tellus*, vol. 47A, p. 974-997.

Toth, Z., et E. Kalnay. 1993. « Ensemble forecasting at NMC: The generation of perturbations ». *Bulletin of the American Meteorological Society*, vol. 74, p. 2317-2330.

_____, et E. Kalnay. 1997. « Ensemble forcasting at NCEP and the breeding method ». *Monthly Weather Review*, vol. 125, p. 3297-3319.

Trevisan, A., et F. Pancotti. 1998. « Periodic Orbits, Lyapunov Vectors, and Singular Vectors in the Lorentz System ». *Journal of the Atmospheric Sciences*, vol. 55, p. 390-398.

Uppala, S. M., P. W. Ilberg, A. J. Simmons, U. Andrae et V. Da Costa Bechtold. 2005. « The ERA-40 reanalysis ». *Quarterly Journal of the Royal Meteorological Society*, vol. 131, 2961–3012.

Vanvyve, E., N. Hall, C. Messager, S. Leroux et J.-P. van Ypersele. 2007. « Internal variability in a regional model over West Africa ». *Climate Dynamics*, vol. 30, p. 191-202 .

Vukicevic, T., et J. Paegle. 1989: « The influence of one-way interacting lateral boundary conditions on predictability of flow in bounded numerical models ». *Monthly Weather Review*, vol. 117, p. 340–350.

Wei, M., Z. Toth, R. Wobus, Y. Zhu, C. H. Bishop et X. Wang. 2006. « Ensemble Transform Kalman Filter-based ensemble perturbations in an operational global prediction system at NCEP ». *Tellus*, vol. 58A, p. 28-44.

Wu, C.-C., J.-H. Chen, S. J. Majumdar, M. S. Peng, C. A. Reynolds, S. D. Aberson, R. Buizza, M. Yamaguchi, S.-G. Chen, T. Nakazawa et K.-H. Chou. 2009a. « Intercomparison of Targeted Observation Guidance for Tropical Cyclones in the Northwestern Pacific ». *Monthly Weather Review*, vol. 137, p. 2471-2492.

Yamaguchi, M., et T. Komori. 2009. « Outline of the Typhoon Ensemble Prediction System at the Japan Meteorological Agency ». *RSMC Tokyo-Typhoon Center Technical Review*, vol. 11, p.

14-24. Available online at http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/techrev.htm, (accessed 2011.2.10).

_____, T. Iriguchi, T. Nakazawa et C.-C. Wu. 2009. « An Observing System Experiment for Typhoon Conson (2004) Using a Singular Vector Method and DOTSTAR Data ». *Monthly Weather Review*, vol. 137, p. 2801-2816.

Yeh, K.-S., J. Côté, S. Gravel, A. Methot, A. Patoine, M. Roch et A. Staniforth. 2002. « The operational CMC-MRB global environmental multiscale (GEM) model. Part III: Non-hydrostatic formulation ». *Monthly Weather Review*, vol. 130, p. 339-356.

Zadra, A., M. Buehner, S. Laroche et J.-F. Mahfouf. 2004. « Impact of the GEM model simplified physics on extratropical singular vectors ». *Quarterly Journal of the Royal Meteorological Society*, vol. 130, p. 2541–2569.

_____, D. Caya, J. Côté, B. Dugas, C. Jones, R. Laprise, K. Winger et L.-P. Caron. 2008. « The Next Canadian Regional Climate Model ». *La Physique au Canada*, vol. 64, no 2, p. 75-83.

_____, M. Roch, S. Laroche et M. Charron. 2003. « The subgrid scale orographic blocking parameterization of the GEM model ». *Atmosphere-Ocean*, vol. 41, p. 155–170.